Regularity of Stochastic Flows of Stochastic Differential Equations with Singular Coefficients and Applications to Finance

DAVID RUIZ BAÑOS

DISSERTATION PRESENTED FOR THE DEGREE
OF PHILOSOPHIAE DOCTOR

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OSLO
2015

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1663

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika Publishing merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Acknowledgements

Først og fremst vil jeg takke Universitetet i Oslo og mine veiledere Prof. Frank Proske, Prof. Giulia Di Nunno og Prof. Bernt Øksendal for å ha tro på meg og gi meg sjansen til å skrive denne avhandlingen. Jeg vil gjerne gi en spesiel takk til min hovedveileder Frank Proske for din uendelige oppmuntring uansett hvor dårlig situasjonen så ut til å være, for din "vi får det til" eller "i verste fall kan vi anta $b = 0$" som hjalp veldig i mange vanskelige omstendigheter som en doktorgrad innebærer. Uten deg og dine vindunderlige idéer ville denne avhandlingen aldri ha sett lyset. For å lære meg at kaffen kan brukes til å skape fantastiske ting, takk Frank nå er jeg blitt avhengig av kaffe!

I would also like to give especial thanks to my second supervisor Giulia Di Nunno for supporting and following me during my four year research and giving me advice in the hard moments. Thank you for your guidance and proximity.

Jeg vil gjerne takke administrasjonen på fakultetet for å gjøre byråkratiet til en mye lettere opplevelse! Spesiel takk til Biljana Dragisic og Elisabeth Seland for støtten og hjelpen.

My co-authors deserve a great special thanks for accompanying me through this long and tough trip as well as my friends and family.

Jeg takker Torstein Nilssen for fire veldig lærerike år med både fritid og forskning. For din tålmodighet og omgjengelighet i våre utallige mange møter om matematikk og felles kopper kaffe med Frank og drøftinger på tavla.

I am also grateful to Prof. Thilo Meyer-Brandis for inviting me over and giving me the opportunity to enjoy such a nice city as Munich in my first research stay abroad, for sharing your great ideas which turned into two papers of this thesis and for many helpful talks we had.

I would like to give heartfelt thanks to my colleague Paul Krühner for uncountably many fruitful discussions and meetings which took form of two papers that made my research really enjoyable. For being a great host during my research stay in beautiful Vienna. With you I found a great collaborator but also a friend. Thank you for your critic view, for your friendship and many chess matches, I am especially thankful for those where I won, not many though, but there are more to come!

També estic molt agraït a Georgina Cerqueda. Per estar al meu costat durant una llarga carrera que m'ha dut fins aquí. Per les moltes vegades que em passat junts i hem sortit, gaudit i viatjat en busca de tresors perduts! Tot això forma part d'aquesta tesi i encara que visquem lluny l'un de l'altre, t'he tingut molt present durant aquests quatre anys.

Thanks to my good friend Krzysztof Paczka. You have been a very important part both in my work and personal life. I would like to thank you for being there. For many dinners and red drinks and long conversations. For your advice and for a fantastic trip in Japan and Barcelona.

A very important thanks to all my colleagues that followed me during my PhD, the list is far from being complete but I would like to mention Hanna Zdanowicz, Elin Røse, Marcus Eriksson, David Martí, Daniel Pérez, Tonje Lien, Erlend Storrøsten, Torquil Sørensen for creating a pleasant environment both in the academic and leisure time.

Mine kollegaer fra den gamle avdeling C for en god motakelse, mange kaffe- og lunsjpauser og flasker vin fra vinlotteriet. Også for felles undervisning som ga meg kompetanse i dette viktige området og ikke minst for å vekke min interesse for statistikk.

A Salvador Ortiz-Latorre i Sara Solanilla per moltes xerrades i riures junts dins i fora la universitat i per donar-me l'oportunitat de poder seguir parlant aquesta llengua tant meravellosa que tenim! Per donar-me suport, pels vostres bons consells i ajuda tant en la recerca com en el personal.

Vull donar especial gràcies a la Prof. Marta Sanz-Solé perquè em vas fer descobrir aquesta branca de les matemàtiques que tant m'agrada. Per els bons consells que em vas donar i la teva ajuda per continuar després del màster. Sense tot això no hauria arribat fins aquí.

Finalment, vull agrair de tot cor la meva família més propera. La meva mare i la meva àvia per haver-me fet qui sóc i cuidat durant tants anys, que em va permetre estudiar i aprendre tot el que sé. Sense elles, tot això hagués estat totalment impossible.

Jeg vil hjertelig takke min norske familie for å ha vært der hele tiden og tatt meg imot med åpne armer. Dere er fantastiske og jeg skylder dere et godt stykke av denne forskingen!

Takk til Simen Østad. Jeg kan knapt uttrykke med ord hvor viktig du har vært for meg i denne langtrukne prosessen. For at du gjorde hverdagen lysere og de vanskelige øyeblikkene lettere. Du er den viktigste medforfatter jeg kunne ha hatt fordi du hjalp meg å takle de mest krevende problemene i en doktorgrad, nemlig de ikke matematiske!

Finally, I would like to thank the committee members, Prof. José Manuel Corcuera, Prof. Arne Løkka and Prof. Tom Lindstrøm for taking the time and effort to evaluate this thesis.

David R. Østad
University of Oslo
August 2015

Contents

Chapter 1

Introduction

1.1 A friendly tour for the layman

Among the diverse fields of mathematics we find *stochastic calculus*, also called *stochastic analysis*, an area which has become more and more interesting to mathematicians and also to scientists due to its numerous applications in real world sciences. The word "stochastic" comes from Greek (mid. 17th century) stokhastikos, from stokhazesthai "aim at, guess, conjecturing". Hence, we can comprehend stochastic calculus as a mathematical tool which aims at "guessing" uncertain phenomena with uncertain outcomes. The core objects within this field are *stochastic processes* which are a collection of *random variables*, that is, variables that incorporate uncertainty or randomness. For instance, if we toss a die, a certain outcome (i.e. with no randomness involved) would be that the die will fall onto the floor, provided that we all agree on the laws of physics. On the other hand, one is unable to predict which face of the die will come up. Nevertheless, we can compute or estimate the probabilities of each outcome. Unfortunately, tossing a die does not fall into the typology of real-life problems a scientist needs to face; for this reason, one needs to develop more advanced machinery to tackle these problems.

Given an experiment of interest, a random variable gives a description of the possible outcomes or events of the experiment. Then we assign a probability to each event, namely, the probability that such event occurs. Now, because some experiments involve several different aspects and also in many cases evolve in time, one is compelled to consider a whole family of random variables. This is known as a stochastic process. Then, one can start studying properties of these objects and build a whole theory on how to handle, interpret and operate these objects and of course, draw conclusions from them.

Some of the vernacular key words associated to the branch of stochastic calculus that one may hear in daily conversation are, for example, experiment, probability, distribution, deterministic, random, uncertainty, stochastic processes, random phenomena, impossible, surely, almost surely, outcome, law, hypothesis, parameter, expectation, mean, variance, volatility, etc.

The interest of stochastic analysis arises in many areas, as in physics, where it is used to explain and model the effects of random motion on physical phenomena. It also occurs in engineering in the so-called filtering problem which, in brief, approaches the problem of trying to find the best estimates of the true value of a system given only some noisy observations of *idem*. Also, in engineering, control theory, which deals with the performance of dynamic systems with inputs and how their behaviour is affected by such inputs. Another stochastic example from finance is that of an agent investing money in the market who changes her strategy according to the random fluctuations of the market prices; here, the way she should make decisions and implement them can be answered using control theory in the stochastic analysis setting. As already hinted, one finds plenty of applications in finance, such as in the theory of pricing and hedging financial derivatives or risk-management which deals with the assessment of risk and its consequences. We find stochastic models in biology for modelling reproduction and environment of populations, as well as in sociology or politics, where one tries to connect theoretical models to the data of sociology; this typically takes the form of surveys performed on individuals or is given as proportions of people doing or believing something. All these examples suggest obtaining an equation or stochastic process based on some theoretical assumptions that try to model the chances of an individual changing state in a given interval of time.

In this thesis we primarily focus on two of the aforementioned topics. First, stochastic calculus itself, involves *stochastic differential equations*, widely known and used in all the preceding areas of application, as well as some of their properties. Secondly, a main concern is the application of stochastic calculus to finance, especially in the so-called *sensitivity analysis of financial options*.

A *stochastic differential equation* can be viewed as a mathematical object which tries to explain certain phenomena in nature, for example. The substantial difference between stochastic differential equations and a classical *differential equation* is the presence of random inputs created by *uncertainty*. In classical analysis an (ordinary) differential equation typically takes the form

$$\frac{\mathrm{d}}{\mathrm{dt}} X_t^x = u(t, X_t^x), \quad X_0^x = x, \quad t \geq 0 \tag{1.1}$$

where here (the input) u gives some information about some experiment or phenomenon and (the output) X represents the total understanding of such phenomenon. Here, X is a "process" or "function", X_t is its associated value at a given time t, (say $t =$'tomorrow') and x denotes the ”initial state”, that is, the state of the system when we start our study, which is, in some cases, completely known. Hence, solving (1.1) (finding X or having full understanding of X) means having full understanding of the phenomenon of study. The equation in (1.1) can in many cases be solved if the "law u" explaining the phenomenon is well behaved—if it behaves quietly and smoothly with no sudden changes or spikes.

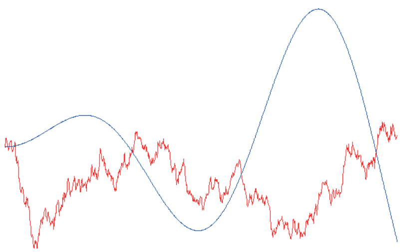

Figure 1.1: Illustration of the difference between *good* (in blue) and *bad* behaviour (in red).

In a summary, one can think of u as "what we know" about an experiment or phenomenon and X of the total answer to understanding why such phenomenon behaves as it does. In classical analysis a model like (1.1) is said to be *deterministic* (as opposed to *stochastic* or *random*), which means that such phenomenon can be predicted perfectly. Examples of this are physical experiments for which well-known laws apply, let us say, dropping a ball from a second floor on a day with no wind. One can by means of (1.1) compute the speed, time and acceleration of the ball at all times until it touches the ground with fair enough accuracy.

There is a whole theory and mathematical machinery to study problems like (1.1) and finding a solution X. Namely, *classical differential analysis* or *differential calculus*, mainly credited to celebrated mathematicians Isaac Newton (1643–1727) and Gottfried Leibniz (1646–1716).

One may already observe that real-life challenges very rarely involve "nice behaviours" and therefore one may fail to find a complete answer X or even any answer at all. Even worse, in real-life experiments, there is a lot of uncertainty involved and u may well be subject to "randomness" and therefore not chosen correctly, which then may lead to wrong predictions. For instance, a widely known example is weather forecasting. Imagine u represents the information available to predict the weather conditions over the next ten days. Then, one finds answer X broadcasted on television and it turns out that we observe different weather. Of course, one concludes that u was not the right "law" to follow.

A possible answer to this problem is to add "randomness" into (1.1), that is

$$\frac{\mathrm{d}}{\mathrm{dt}}X_t^x = u(t, X_t^x) + \text{"noise"}, \quad X_0^x = x, \quad t \geq 0 \tag{1.2}$$

where, indeed, the term "noise" introduces *uncertainty* to our predictions and studies of a problem. There is an enormous variety of ways to introduce noise into a model. One of the most celebrated noises is the so-called *Wiener process* which describes *Brownian motion*. Since this thesis is notably based on and uses Brownian motion as driving noise for the modelling of uncertainty we should not fail to mention the findings of Robert Brown (1773–1858), a Scottish botanist and palaeobotanist who observed what is today called Brownian motion in the movement of pollen grains of the plant *Clarkia pulchella* suspended in water under a microscope.

... While examining the form of these particles immersed in water, I observed many of them very evidently in motion; their motion consisting not only of a change of place in the fluid, manifested by alterations in their relative positions, but also not unfrequently of a change of form in the particle itself; a contraction or curvature taking place repeatedly about the middle of one side, accompanied by a corresponding swelling or convexity on the opposite side of the particle. In a few instances the particle was seen to turn on its longer axis. These motions were such as to satisfy me, after frequently repeated observation, that they arose neither from currents in the fluid, nor from its gradual evaporation, but belonged to the particle itself. Fragment of the original document by Robert Brown (1828), Phil. Mag. 4, 161–173.

In the context of mathematics, the "random jittery" motion is modelled by the red curve in Figure 1.1. Using the standard denotation for Brownian motion "$\frac{\mathrm{d}B_t}{\mathrm{d}t}$", the equation (1.2) now reads as

$$\frac{\mathrm{d}}{\mathrm{dt}}X_t^x = u(t, X_t^x) + \text{"}\frac{\mathrm{d}B_t}{\mathrm{d}t}\text{"}, \quad X_0^x = x, \quad t \geq 0. \tag{1.3}$$

A solution to problem (1.3) is typically a stochastic process, i.e. a family of random variables, hence, the uncertainty is captured at once by X, saying that, for a given time t, the state of the nature X_t is a random variable, i.e., not completely determined. Some examples where equation (1.3) is used are in modelling the fluctuations of prices in the stock market, the movement of particles in a fluid due to collisions with fluid molecules (so-called Langevin equation), weather prediction, etc.

As it was the case for the (deterministic) problem (1.2), here one also needs to develop a (stochastic) *differential calculus* to deal with objects like (1.3) and to solve them. The most widely used calculus to treat (1.3) is *Itô calculus*, attributed to Japanese mathematician Kiyoshi Itô (1915–2008). He defined the concepts of *Itô stochastic process*, *Itô integral* and a whole methodology of rules to deal with stochastic differential equations and how to solve them.

It is very worth mentioning that the classical problem described in (1.1) where no randomness is included may fail to have a unique answer (solution) or, even worse, it may fail to have any answer at all. This is the case if, for instance, the object u behaves like the red curve in Figure 1.1 instead of the blue one. One of the most remarkable peculiarities about stochastic differential equations is that the corresponding equivalent problem in (1.2) admits a unique solution X even if the input u behaves roughly.

In summary, before we proceed to greater technicalities, this thesis deals with the study of objects like (1.3) when u is very irregular and how to construct a solution (answer) X in that case. Also, we give an example where this can be applied in the context of finance. We also study the problem of changing the *noise* "$\frac{\mathrm{d}B_t}{\mathrm{d}t}$" into much *rougher* noise, that is, much more irregular than the red curve in Figure 1.1 in order to see what kind of impact this has on solution X.

1.2 Context and lines of investigation

This thesis, presented for the degree of Doctor Philosophiae, consists of six articles written over the last four years. The intention of this section is to embed the thesis into its mathematical framework, describe the different lines of investigation taken into consideration and give a brief résumé of the state of the field.

The common line of research in all of these articles is to provide a better understanding of the behaviour and effects of *Brownian motion* in differential equations. We consider *stochastic differential equations* driven by Brownian motion, or *fractional* Brownian motion in the last article, as the core object of study, that is

$$dX_t^x = b(t, X_t^x)dt + dB_t, \quad X_0^x = x \in \mathbb{R}^d, \quad t \in [0, T], \tag{1.4}$$

where $\mathbb{R}^d$ denotes the d-dimensional Euclidean space of real numbers, $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector field indexed by t, $B = \{B_t, t \in [0, T]\}$ is a Brownian motion relative to a stochastic basis

$$(\Omega, \mathcal{F}, P), \quad \{\mathcal{F}_t\}_{t\in[0,T]}$$

where Ω is the sample space, $\mathcal{F}$ the σ-algebra of events, P a probability on $(\Omega, \mathcal{F})$ and $\{\mathcal{F}_t\}_{t\in[0,T]}$ the family of σ-algebras generated by the random variables B_t, $t \in [0, T]$ including all P-null sets. Finally, $x \in \mathbb{R}^d$ is the initial state which is taken to be deterministic and T is the final time horizon.

The stochastic process B is defined as the process satisfying the following four conditions

1. $B_0 = 0$, P-a.s.
2. The increments of B are independent, i.e., $B_{t_4} - B_{t_3}$ and $B_{t_2} - B_{t_1}$ with times $t_4 > t_3 > t_2 > t_1$ are independent random variables.
3. The increments are stationary, i.e., given times $t_2 > t_1$ the law of $B_{t_2} - B_{t_1}$ is the same as the law of $B_{t_2-t_1}$.
4. At each given time $t \geq 0$, the law of the random variable B_t is normally distributed with zero mean and variance t, i.e., in notation form, $B_t \sim N(0, t)$.

It follows from Kolmogorov's continuity criterion that there is a version with *almost surely* continuous sample paths.

If problem (1.4) admits a solution $X = \{X_t, t \in [0, T]\}$ it makes sense to write

$$X_t = x + \int_0^t b(s, X_s^x)ds + B_t \tag{1.5}$$

where the integration is understood in the sense of Lebesgue, hence we will just consider vector fields b for which the integral makes sense. For example, if b is Lipschitz continuous uniformly in time t and has, at most, linear growth, then one has the classical *existence and uniqueness* result for SDEs like (1.4). Unfortunately, in many applications where SDEs appear, as mentioned in the previous section, the vector field b is far from being Lipschitz and the natural question arises of whether a solution exists or not, for instance, if b is discontinuous.

1.2.1 Existence and uniqueness of solutions

The interest in studying such equations can be justified by looking at the classical Cauchy problem

$$dX_t^x = b(t, X_t^x)dt, \quad X_0^x = x \in \mathbb{R}^d, \quad t \in [0, T], \tag{1.6}$$

where we have removed the source of *noise* B. Now, if b is "nice" enough we know, by classical results, that (1.6) admits a unique solution X. For instance if b is Lipschitz continuous, i.e., there is a finite constant $C > 0$ independent of t such that

$$|b(t, x) - b(t, x)| \leq C|x - y|, \quad x, y \in \mathbb{R}^d,$$

then Picard-Lindelöf's theorem guarantees (local) existence and uniqueness of a differentiable function $[0, T] \ni t \mapsto X_t \in \mathbb{R}^d$ such that (1.6) holds. Nevertheless, if b fails to be Lipschitz, a solution to problem (1.6) might fail to be unique, or even to exist. For example, if $x = 0$ and $b(t, y) = y^{2/3}$, which is not Lipschitz continuous, then uniqueness breaks down. Both $X_t = 0$ and $X_t = t^3$ solve the equation. A remarkable effect of adding Brownian noise into (1.6) is that it possesses a "regularising" effect on (1.6) in the sense that, even if b is not Lipschitz, existence and uniqueness of solutions of (1.4) may still be fulfilled even if b has discontinuities.

For studying existence recall that one may distinguish between two different concepts of solutions of (1.4). This means it is sometimes easier to prove that relation (1.4) holds for two given pair of processes (X, B) by a simple application of Girsanov's theorem but it is not always the case that X is adapted to the filtration of B, i.e., X can not always be represented as a functional of the driving noise. In this case we say a pair (X, B) satisfying (1.4) where X need not be adapted to $\{\mathcal{F}_t\}_{t\in[0,T]}$ is a *weak solution* of (1.4). On the contrary, if X is adapted then we say X is a *strong solution*.

For studying uniqueness, one may discriminate between different concepts of uniqueness; we say two solutions are *weakly unique* or *unique in law* if their finite dimensional laws coincide. On the other hand, we say *strong uniqueness* or *pathwise uniqueness* holds if the solutions defined on the same probability space agree on a full-measure set. Of course, pathwise uniqueness implies uniqueness in law.

In 1965, A. V. Skorokhod in [104] showed that there are solutions of SDEs under the condition that the coefficients are just continuous; for this reason, the problem of uniqueness becomes important. In 1971 T. Yamada and S. Watanabe, in [108], contributed with a new method to prove existence and uniqueness of strong solutions of SDEs, relying on the fact that two weak solutions that are pathwise unique must be strong, in short, pathwise uniqueness implies strong existence. A breakthrough within the theory of stochastic differential equations was to prove that problem (1.4) admits a unique strong solution even if b is merely measurable and bounded. This was done by A. K. Zvonkin in [110] in the one-dimensional case in 1974 and then generalised to several dimensions five years later by A. Y. Veretennikov in [105]. Both authors employ the Yamada-Watanabe principle in connection with Kolmogorov's equation.

Other examples of the construction of solutions of SDEs with irregular coefficients based on a pathwise uniqueness argument are N. V. Krylov and M. Röckner in [68], where the drift coefficient is assumed to be integrable. Other examples are [54] and [55], as well as in [28] and [29] in infinite dimensions. A widely used approach is what the authors in [29] call the "Itô-Tanaka trick" which makes use of the (backward) Kolmogorov's equation associated with the diffusion (1.4). Consider the parabolic partial differential equation

$$\partial_t u + b\nabla u + \frac{1}{2}\Delta u = b \tag{1.7}$$

on $[0, T] \times \mathbb{R}^d$. One may replace the irregular term in (1.5) by the following expression by simply using Itô's formula on $u(t, X_t)$. Namely,

$$\int_0^t b(s, X_s)ds = u(t, X_t) - u(0, x) - \int_0^t \nabla u(s, X_s)dB_s, \quad t \in [0, T], \tag{1.8}$$

where now u solves (1.7) and is better behaved. Certainly, one of the limitations of this approach is that the system must be Markovian, which, for instance, we do not assume in a case in which we consider a fractional noise instead of the noise B.

To overcome this limitation we employ a different approach based on the so-called Malliavin calculus. Before we go into the details of the method we will first outline a short overview of this topic.

1.2.2 Malliavin's calculus of variations

Another major tool exploited in this thesis is *Malliavin calculus*. This type of calculus in some sense extends the classical variational calculus for functions to stochastic processes. It is a variational calculus and allows for the computations of derivatives of random variables in a certain sense.

This calculus is attributed to the French mathematician Paul Malliavin. His motivation came from Hörmander's results on a sufficient condition, the so-called Hörmander's condition, for a differential operator to be *hypoelliptic*. Malliavin wanted to give a probabilistic proof that Hörmander's condition implied that the density of a solution of a stochastic differential equation is smooth, see [78], while Hörmander's proof was based on the theory of partial differential equations. For this purpose, he developed a variational calculus on the Wiener space which allows for "differentiation" of random variables in a certain sense. His calculus enables him to prove properties for the densities of random variables and in particular study regularity properties of the densities of solutions of SDEs and to find bounds. In fact, Malliavin's calculus of variations has had an even greater impact than just studying densities of random variables. For instance, one can construct an anticipative stochastic calculus using Malliavin calculus which has abundant applications in finance, for example in cases of insider information. Also, Malliavin calculus allows one to determine explicitly the kernel in the stochastic integral of the *martingale representation theorem*, i.e., the so-called Clark-Ocone formula.

The idea of how Malliavin calculus was developed can be stated in a simple way, as an extension of classical calculus. It is known that the Lebesgue measure has the following invariance property on $\mathbb{R}$, that is, for any integrable function f in the sense of Lebesgue and any real number $\varepsilon > 0$ we have

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} f(x+\varepsilon)dx.$$

The latter allows us to derive the well-known integration by parts formula by simply choosing $f = gh$ where g and h are two integrable functions, then differentiating with respect to ε, i.e.

$$\int_{\mathbb{R}} f'(x+\varepsilon)dx = \int_{\mathbb{R}} g'(x+\varepsilon)h(x+\varepsilon)dx + \int_{\mathbb{R}} g(x+\varepsilon)h'(x+\varepsilon)dx$$

and finally using the invariance property

$$\int_{\mathbb{R}} (gh)'(x)dx = \int_{\mathbb{R}} g'(x)h(x)dx + \int_{\mathbb{R}} g(x)h'(x)dx.$$

The aim is to extend this idea to random variables in a probability space $(\Omega, \mathcal{F}, P)$ where $\Omega = C_0([0,1])$ is the Wiener space and P is the Wiener measure. It is known that the Wiener measure does not satisfy the invariance property as is the case of the Lebesgue measure on $\mathbb{R}$. Nevertheless, the Cameron-Martin theorem gives the corresponding factor which arises from "translating" P. If W is a Wiener process and h is a square integrable predictable process then in the case of random variables one has

$$E\left[F\left(W+\varepsilon\int_0^t h_s ds\right)\right] = E\left[F(W)\exp\left\{\varepsilon\int_0^1 h_s dW_s - \frac{1}{2}\varepsilon^2\int_0^1 h_s^2 ds\right\}\right].$$

Then differentiating with respect to ε on both sides and evaluating at $\varepsilon = 0$ we obtain the following *integration by parts formula*, also known as *duality relation*

$$E\left[\langle DF(W), \int_0^t h_s ds\rangle\right] = E\left[F(W)\int_0^1 h_s dW_s\right]$$

where the left-hand side represents the Malliavin derivative of the random variable F in the Cameron-Martin direction $\int_0^t h_s ds$. The stochastic integral on the right-hand side is in the sense of Itô. The expression still makes sense even for non-adapted integrands h as long as the integration is understood in the sense of Skorokhod.

Malliavin calculus has also been further developed by Srinivasa S.R. Varadhan, Daniel Stroock, Jean-Michel Bismut, Shinzo Watanabe, David Nualart, Denis Bell and others.

Nowadays, there are two common approaches to Malliavin calculus and the Malliavin derivative. One is based on the Wiener-Itô chaos decomposition which basically states that any random variable in $L^2(\Omega)$ can be represented as a series of iterated Itô integrals, i.e.

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad F \in L^2(\Omega)$$

where I_n represents an iterated Itô integral, that is

$$I_n(f_n) = \int_{[0,T]^n} f_n(t_1,\ldots,t_n)dW_{t_1}\cdots dW_{t_n}$$

for a suitable deterministic symmetric kernel $f_n \in L^2([0,T]^n)$, $n \geq 0$. The convergence of the above series is in $L^2(\Omega)$ and one has the corresponding Itô isometry

$$\|F\|^2_{L^2(\Omega)} = \sum_{n=0}^{\infty} n!\|f_n\|^2_{L^2([0,T]^n)}.$$

Then we say that F is Malliavin differentiable and denote the space of Malliavin differentiable random variables by $\mathbb{D}^{1,2}$, if

$$\|F\|^2_{\mathbb{D}^{1,2}} := \sum_{n=0}^{\infty} nn!\|f_n\|^2_{L^2([0,T]^n)} < \infty.$$

If the above sum converges then $F \in \mathbb{D}^{1,2}$ and we define the Malliavin derivative D_tF of F at time t as the expansion

$$D_tF = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot,t)), \quad t \in [0,T],$$

where $I_{n-1}(f_n(\cdot,t))$ is the $(n-1)$-fold iterated integral of $f_n(t_1,\ldots,t_{n-1},t)$ with respect to the first $n-1$ variables and $t_n = t$ is left as parameter. See [36] for more details.

The second common approach to Malliavin calculus is built in more general terms via a closable operator defined on a space of simple random variables: Let $W = \{W(h), h \in H\}$ an *isonormal Gaussian process* associated with a Hilbert space H with scalar product $\langle\cdot,\cdot\rangle$. Assume W is defined on a complete probability space $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is generated by W. Let $\mathcal{S}$ denote the space of *smooth random variables* in the sense that $F \in \mathcal{S}$ has the form

$$F = f(W(h_1),\ldots,W(h_n)), \quad h_1,\ldots,h_n \in H, \quad f \in C_0^\infty(\mathbb{R}^n).$$

Then the Malliavin derivative of a smooth random variable is defined as the H-valued random variable

$$DF = \sum_{i=1}^n \partial_i f(W(h_1),\ldots,W(h_n))h_i.$$

It can be proven that D is a closable operator from $L^2(\Omega)$ to $L^2(\Omega;H)$. The domain of D is usually denoted by $\mathbb{D}^{1,2}$, which coincides with the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,2} := \|F\|_{L^2(\Omega)} + \|DF\|_{L^2(\Omega;H)}.$$

One can also verify that D satisfies most common properties, such as the product rule, the chain rule, an integration by parts formula, etc.

In this thesis we mainly treat the case when $H = L^2([0,T])$ with $T > 0$ being a finite time horizon and $W(h) = \int_0^T h_s dW_s$ is the usual Itô integral.

To end this section, we present what we consider to be a very important result in the context of Malliavin calculus, specially when applied to stochastic differential equations: a compactness criterion for subsets of $L^2(\Omega)$ due to Giuseppe Da Prato, Paul Malliavin and David Nualart which can be found in [30]. Essentially, the criterion states that if one can control the Malliavin derivatives of a sequence of random variables in $L^2(\Omega)$ and the Malliavin derivatives possess some Hölder-regularity, then the sequence is relatively compact. In other terms, if $\{X_n\}_{n\geq 0} \subset L^2(\Omega)$ and

$$\sup_{n\geq 0} \|X_n\|_{1,2} = \sup_{n\geq 0} \|X_n\|_{L^2(\Omega)} + \sup_{n\geq 0} \|D.X_n\|_{L^2([0,T]\times\Omega)} < \infty$$

and

$$\int_0^T \int_0^T \frac{E\left[|D_s X_n - D_{s'} X_n|^2\right]}{|s-s'|^{1+2\beta}} ds ds' < \infty$$

for some $\beta > 0$ then the sequence $\{X_n\}_{n\geq 0}$ is relatively compact in $L^2(\Omega)$. This criterion will be used in Chapter 5, Chapter 6 and Chapter 7 to construct solutions to SDEs where the random variables X_n, $n \geq 0$ will play the role of an approximating sequence of the solution.

1.2.3 A new method to construct strong solutions based on the Malliavin calculus of variations

In recent years, Frank Proske and Thilo Meyer-Brandis in [83] have developed a new method for constructing (unique) strong solutions of SDEs for irregular coefficients based on this new variational calculus for stochastic processes. The novelty of the method is that not only does it not rely on the Yamada-Watanabe principle but it also gives the additional insight that such solutions are differentiable in the Malliavin sense. Thus, the method indicated that the property of a solution being Malliavin differentiable is solidly linked to the "nature" of strong solutions. This method is based on the compactness criterion mentioned at the end of the previous section and it is used to show that the solution can be approximated by a sequence of processes that are compact in the linear subspace of adapted processes in $L^2(\Omega)$. Finally, uniqueness in law is enough to verify strong uniqueness. Hence, this method is in some sense opposed to the Yamada-Watanabe principle. Here, one starts with proving strong existence and then uniqueness in law. The Malliavin differentiability then follows automatically since one is compelled to prove that the Malliavin norms are uniformly bounded. This method is also applied in [81], in [87] to construct Sobolev differentiable stochastic flows associated with a solution of (1.4) and in [88] to construct solutions with even discontinuous unbounded drift coefficients in the one-dimensional case.

Recall the following stochastic differential equation in the setting of (1.4)

$$dX_t^x = b(t, X_t^x)dt + dB_t, \quad X_0^x = x \in \mathbb{R}^d, \quad t \in [0, T]$$

where b is a very irregular vector field.

This method can mainly be divided into the following central steps:

- One approximates the irregular term b by a sequence of nicely-behaving functions $\{b_n\}_{n\geq 0}$. Then X_t^n denotes the solution of the SDE when we replace b with b_n.

- One constructs a weak solution (X, B), usually by means of Girsanov's theorem. *A priori*, X does not need to be adapted to the filtration generated by B.

- One shows that the sequence of well-behaved solutions X_t^n converges to $E[X_t|\mathcal{F}_t]$ in the weak topology of $L^2(\Omega)$.

- By use of the compactness criterion from [30] as explained at the end of the previous section, one can show that $\{X_t^n\}_{n\geq 0} \subset L^2(\Omega)$ is relatively compact. Hence, by the previous step one has that X_t^n converges to $E[X_t|\mathcal{F}_t]$ in $L^2(\Omega)$ and that $E[X_t|\mathcal{F}_t]$ is Malliavin differentiable.

- A transformation property of the type $E[\varphi(X_t)|\mathcal{F}_t] = \varphi(E[X_t|\mathcal{F}_t])$ for all bounded continuous functions φ allows us to conclude that $E[X_t|\mathcal{F}_t] = X_t$ being thus X adapted to $\{\mathcal{F}_t\}_{t\in[0,T]}$, and hence is a strong solution.

- Finally, we prove that the solutions are unique in law and since they are strong then pathwise uniqueness holds.

This method is very general and can be applied to a wide class of SDEs. The reason that the SDE has additive noise simplifies the computations considerably. Nevertheless, one can consider more general C^1 bounded diffusions, or, by a simple application of Itô's lemma one can include a large class of non-trivial diffusion coefficients which are found in many applications. Another main advantage of this method is that, for instance, no PDE theory is required and no

Markovianity of the system is assumed. Hence, while the "Itô-Tanaka" trick in (1.8) may fail, as is the case of fractional Brownian motion as driving noise, the method described above can still be employed. Chapter 7 details an example of this. Last but not least, the method can also be applied to even infinite dimensional non-Markovian systems as for instance

$$dX_t = (AX_t + b(X_t))dt + QdW_t^H, \quad X_0 = x \in H$$

for mild solutions X, where A is a densely defined linear operator on a separable Hilbert space H, $b : H \to H$ is an irregular functional, Q a Hilbert-Schmidt operator and W^H a cylindrical Brownian motion.

1.2.4 Regularity of the solution and the Bismut-Elworthy-Li formula

One important direct application we find of the Malliavin differentiability is the derivation of the so-called Bismut-Elworthy-Li formula. This formula is a representation of the spatial derivatives of the solution of Kolmogorov's equation; in other words, it gives an expression for the derivative in space of a strongly continuous semigroup associated to a Markovian diffusion independently of the derivative of the function involved. This formula was shown in [21] and further extended in [40]. In [50] the authors use a different technique relying on Malliavin calculus to prove the formula and then use it for the computation of Δ-sensitivities of financial options. As mentioned, the additional Malliavin regularity can help derive the Bismut-Elworthy-Li formula, which is an important application within finance. Hence, when constructing Malliavin differentiable strong solutions via the method described in the previous section, one can directly derive the corresponding Bismut-Elworthy-Li identity associated with the strong solution. These two aspects, the construction of solutions and the corresponding Bismut-Elworthy-Li formula, are central common characteristic of this thesis.

Another important feature in the study of SDEs and their regularity is the study of the densities of their solutions. There is very little known about the densities of solutions of SDEs with very irregular coefficients. Nevertheless, some advances in this direction have been made. For instance, M. Hayashi, A. Kohatsu-Higa and G. Yûki in [57] showed that SDEs with Hölder continuous drift and smooth elliptic diffusion coefficients have solutions with Hölder continuous densities at any time. An important tool for studying regularity of densities are *integration by parts formulas*. V. Bally and L. Caramellino in [5] derive an integration by parts formula and relate the integrability properties of the weight to the regularity of the density of the underying process. Also, S. De Marco in [31] proved smoothness of the density on an open domain where the usual conditions of ellipticity and smooth coefficients on such domain are fulfilled. A remarkable fact is that Hörmander's condition is not used. Both results rely substantially on Malliavin calculus and the aforementioned integration by parts formula in connection with tail estimates on the Fourier transform of the solution. One must also mention the results of V. Bally and A. Kohatsu-Higa [6], in which they provide bounds for the density of a type of a two-dimensional degenerated SDE. For this case, it is assumed that the coefficients are five times differentiable with bounded derivatives. A curious result in the same direction is by A. Kohatsu-Higa and A. Makhlouf in [64] where they show smoothness of the density for smooth coefficients that may also depend on an external process whose drift coefficient is irregular.

Upper and lower estimates for the density are also given.

As one can observe it seems very demanding to obtain further *good* properties of densities for very *bad* coefficients. Malliavin calculus appears to be a widely common tool for dealing with problems involving study of densities. Nevertheless, it seems troublesome to gain regularity with singular coefficients via Malliavin calculus. For example, let us point out a notable result by A. Debussche and N. Fournier in [32] on this topic, in which they show that the finite dimensional densities of a solution of an SDE with jumps lies in a certain (low regular) Besov space when the drift is Hölder continuous. The novelty is that their approach does not rely on Malliavin calculus. To give an idea why we believe Malliavin calculus might not be the best tool to study densities of solutions of SDEs with very irregular coefficients one can look at the result in Chapter 2 where it is shown that an SDE with Lipschitz continuous drift has a twice Malliavin differentiable solution and for this reason its density is C^{α}-Hölder continuous for any $\alpha \in (0,1)$. On the contrary, the SDE (1.4) with drift coefficient $b(t,x) = 1_{\{x>0\}}$ which is discontinuous has a Lipschitz continuous density. One can even construct a random variable in $\mathbb{D}^{1,1}$ which has a continuous density or which does not admit a density at all.

In this thesis we are also concerned with this problem as we believe it improves one's understanding of the nature of SDEs driven by Brownian motion and their regularity. Our results in this direction have been to develop a new method for studying densities of Itô-type processes and obtain very explicit lower and upper bounds for the densities in a very general (even non-Markovian) context which, as an innovation, does not rest on Malliavin calculus techniques.

In a summarised picture, the idea is the following. Consider the following family of processes

$$dX_u(t) = u(t)dt + dW(t), \quad X_u(0) = 0 \in \mathbb{R}^d, \quad t \in [0,T], \tag{1.9}$$

where W is a given d-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ generated by $W(t), t \geq 0$. The process u is a bounded and $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted process with integrable trajectories. We are only concerned with weak solutions so u bounded is enough to guarantee existence of a weak solution of (1.9) which admits a density ρ_t for every $t > 0$. Then one can show that the density of $X_u(t)$ can be computed as

$$\rho_t(x) = \limsup_{\varepsilon\to 0} \frac{P(|X_u(t) - x| \leq \varepsilon)}{V_\varepsilon},$$

where $V_\varepsilon = \frac{\pi^{d/2}\varepsilon^d}{\Gamma(d/2+1)}$ is the Lebesgue measure of the d-dimensional Euclidean ball with radius $\varepsilon > 0$ and Γ here denotes the *gamma* function. Because of the above expression, studying bounds for the density ρ_t can be reduced to studying bounds for the distribution function $P(|X_u(t) - x| \leq \varepsilon)$. In other words, to obtain the upper-bounds for ρ_t we need to obtain the biggest values of $P(|X_u(t) - x| \leq \varepsilon)$. This can be posed as a control problem in the following way

$$\sup_{u\in\mathcal{A}} P(|X_u(t) - x| \leq \varepsilon),$$

where $\mathcal{A}$ denotes the set of admissible controls u, i.e., the set of bounded and adapted processes. Intuitively, the process that maximises the probability of being near 0, that is, in $[-\varepsilon, \varepsilon]$ is $u(x) = -\text{sign}\,(x)$ where sign denotes the generalised signum function, i.e., $\text{sign}\,(x) =$

$\frac{x}{|x|}1_{\{x\neq 0\}}$, $x \in \mathbb{R}^d$. Similarly, the process that minimises this probability is then $u(x) = \text{sign }(x)$. Hence, the densities of the solutions $X_{\text{-sign}}(t)$ and $X_{\text{ sign}}(t)$ provide upper- and lower-bounds for the densities of any process of the form $X_u(t) = x + \int_0^t u(s)ds + W(t)$, $u \in \mathcal{A}$, $t \in [0, T]$. Observe that u here is very general so in particular this class of processes include solutions of SDEs with merely bounded drift and are possibly path-dependent as well. We believe this method can be further studied in detail to obtain more regularity of the densities although this may be a difficult task. A work in progress on this issue in [13] revealed that the densities are in fact Hölder continuous of any order $\alpha \in (0, 1)$ in dimension one. Indeed, by employing the same idea, the control problem is now considered on the Fourier-Stiltjes transform of the law of $X(t)$, $t > 0$. Thus, finding the *worst* characteristic function among all characteristic functions of SDEs with bounded, measurable and path-dependent drift coefficient. As a consequence, we show that the fundamental solution of the Fokker-Planck equation in dimension one is even Hölder continuous of order $\alpha \in (0, 1)$ which remained as an open question.

1.2.5 Mathematical finance

Another prominent line of investigation is *finance* or rather, *mathematical finance*, which can be regarded as the field of applied mathematics concerned with financial markets. In general, a financial mathematician is concerned with modelling financial asset values, as for instance, the value of a commodity or financial asset, the price of a company's shares, etc. While economists may try to find an explanation or reason why a company has a certain share price, a financial mathematician will take the price as given and will try to derive, by means of formulas and models from stochastic calculus, values for the so-called financial derivatives and contracts written on the stock of interest. One can mainly divide this area into two main categories: on the one hand, *derivative pricing theory* and on the other, *risk- and portfolio management*. In both one tries to assess certain quantities and assign a value to them when one has a given model as underlying dynamics. Then the so-called *sensitivity analysis* plays a role. Usually, models in stochastic analysis depend on unknown parameters that have to be estimated, as, for example, drift parameters, volatility, initial value, interest rate parameters, etc. One then wishes to derive the price of a contract and compute how "sensitive" this price is with respect to variations in the underlying parameters. The latter concerns us in one of the scientific articles. More concretely, we show that for an SDE with irregular coefficient, one can still study the sensitivity of the solution with respect to the initial condition in a classical sense. For this purpose we make use of Malliavin calculus and techniques based on integration with respect to the local time of a process. As mentioned, we also derive in this case a Bismut-Elworthy-Li formula for the derivative with respect to the initial condition of the price of an option, independently of the derivatives of the functions involved in the model. The first authors to use Malliavin techniques to study sensitivities were E. Fournié, J-M. Lasry, J. Lebuchoux, P-L. Lions and N. Touzi in [50]. Our results can be regarded as a non-trivial extension of the latter to the case of discontinuous and unbounded coefficients both in the drift and pay-off function in dimension one, as well as some proposed approximations to treat Asian-type options which usually involve Skorokhod-type integrals that are in most cases hard to simulate.

1.2.6 Regularising effects of fractional Brownian noise

Finally, we go one step further by looking at *fractional Brownian motion* which is a generalisation of Brownian motion. Fractional Brownian motion is a Gaussian stationary process which was first introduced by Benoit B. Mandelbrot and John W. Van Ness in the paper *Fractional Brownian motions, fractional noises and applications* in 1968, see [80]. The following is the exact definition given in their paper: let $0 < H < 1$ and let b_0 be an arbitrary number. We call the following random function $B_H(t,\omega)$ *fractional Brownian motion* with parameter H and starting value b_0 at time $t = 0$. For $t > 0$, $B_H(t,\omega)$ is defined by

$$\begin{aligned} B_H(0,\omega) &= b_0 \\ B_H(t,\omega) - B_H(0,\omega) &= \frac{1}{\Gamma(H+1/2)} \left\{ \int_{-\infty}^{0} \left[(t-s)^{H-1/2} - (-s)^{H-1/2} \right] dB(s,\omega) \right. \\ &\qquad \left. + \int_0^t (t-s)^{H-1/2} dB(s,\omega) \right\}, \end{aligned}$$

where the stochastic integration is understood in both the pathwise sense and L^2-sense.

The authors explain that fractional Brownian motion can be used to describe numerous phenomena in nature, for instance in the study of fluctuations in solids. Another class of phenomena with extremely long dependence is encountered in hydrology: Hurst 1951, 1956 discovered that the range of cumulated water flows changes proportionately to dH with $1/2 < H < 1$. Hurst's law has significant practical importance in the design of water systems.

Another application which is more related to our research interest is within economics and is explicitly stated in the paper as follows

It is known that economic time series "typically" exhibit cycles of all orders of magnitude; the slowest cycles have periods of duration comparable to the total sample size. The sample spectra of such series show no sharp "pure period" but a spectral density with a sharp peak near frequencies close to the inverse of the sample size. (B. Mandelbrot, J. W. Van Ness. SIAM Review: 10 (4), 1968, 422-437)

Patrick Cheridito has studied the use of fractional Brownian motion in finance. He found out in [26] that models involving this process may give rise to the presence of a weak form of arbitrage, the so-called "free lunch with vanishing risk" introduced by Freddy Delbaen and Walter Schachermayer in [34] due to the fact that B_H is not a semimartingale whenever $H \neq 1/2$. Nevertheless, he manages to rule out arbitrage strategies by introducing a minimal amount of time that lies between two consecutive transactions which, on the other hand, is a very plausible assumption.

Fractional Brownian motion can also be defined as a process $B_t^H, t \in [0,T]$ on a probability space $(\Omega, \mathcal{F}, P)$ that is centred Gaussian with covariance function given by

$$R_H(t,s) := E[B_t^H B_s^H] = \frac{1}{2}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right), \quad t,s \geq 0.$$

The process B^H enjoys the following self-similarity property

$$\{B^H_{\alpha t}\}_{t\geq 0} \overset{law}{=} \{\alpha^H B^H_t\}_{t\geq 0}$$

for all $\alpha \geq 0$. In fact, fractional Brownian motion is the only stationary Gaussian process satisfying the latter property.

Now then, it is interesting to study the following stochastic differential equation with fractional noise

$$X_t^x = x + \int_0^t b(X_s^x)ds + B_t^H, \quad 0 \leq t \leq T, \quad x \in \mathbb{R}^d \tag{1.10}$$

where $B^H = \{B_t^H, t \in [0,T]\}$ is a d-dimensional fractional Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}_{t\in[0,T]}$ where the filtration is generated by B^H and augmented by all P-null sets.

The main difficulties that one faces in this context are that B^H does not satisfy the Markov property and hence, the increments are not independent. Another difficulty encountered is that B^H, $H \neq 1/2$ is not a *weak semimartingale* as mentioned before. For these reasons it becomes an arduous task to construct a stochastic calculus based on B^H, $H \neq 1/2$, especially in the case $H < 1/2$. No well-established Itô's formula is known for general Itô processes and very little is known about its corresponding Kolmogorov's equation. For this reason, constructing strong solutions to SDE's with even additive fractional noise as we explained for the case $H = 1/2$ in (1.8) is no longer possible as no PDE theory for B^H is established for $H < 1/2$.

Some results in one dimensional have been achieved by using comparison theorems. For example, in 2002 David Nualart and Youssef Ouknine in [91] prove that there is a unique strong solution of SDE (1.10) wirh $H < 1/2$ when b is bounded and measurable and $d = 1$. The method is based on the Yamada-Watanabe theorem and comparison theorems and hence the dimensional restriction.

In Chaper 7, using the method based on Malliavin calculus introduced before we manage to overcome the constraints encountered by considering fractional noise with small Hurst parameters and thus are able to construct unique strong solutions of (1.10) with singular drift coefficients, even in high dimensions, for the first time. We see this case as an example of the strength of the method to construct strong solutions of SDEs. Furthermore, we show that B^H possesses a regularising effect on the solution X seen as a function of the initial value, i.e. $x \mapsto X_t^x$. Namely, we are able to show that the rougher the noise B^H is, the more regular $x \mapsto X_t^x$ gets. In other words, we show that for a small enough Hurst parameter $H < H(k)$, being k some natural number, we have

$$\{x \mapsto X_t^x\} \in L^2\left(\Omega, C^k(\mathbb{R}^d)\right).$$

1.3 Structure of the thesis and contributions

This thesis can be divided into four main parts, all of which explore stochastic differential equations, their properties and some applications. These four parts can be embedded into two main groups: nicely-behaving coefficients in the two first chapters and irregular coefficients in the left chapters. In the first stage we have Chapter 2 and Chapter 3 dealing with the regularity of solutions in the Malliavin and Sobolev sense and the Bismut-Elworthy-Li formula for mean-field SDEs. Chapter 4 deals with the problem of studying densities of Itô-type processes and finding lower- and upper-bounds. For this purpose a new method is presented which in some sense overcomes the limitations of Malliavin calculus for the study of similar problems. This method can further be investigated as a future research. Chapter 5 and Chapther 6 are related to each other. In both we construct strong solutions of SDEs when the drift coefficient is irregular and derive the corresponding Bismut-Elrworthy-Li formula in each case. Finally, in Chapter 7 we start with the study of SDEs driven by fractional noise. This includes the first steps toward future research in this direction.

Chapter 2 is extracted from the paper *Malliavin and flow regularity of SDEs. Application to the study of densities and the stochastic transport equation*, [14] co-authored with T. Nilssen. We give a criterion to determine the regularity of solutions of SDEs in both the Malliavin and Sobolev sense according to how regular the drift coefficient is. As an application we can improve the regularity of densities of such solutions. Also, as a direct consequence, we can prove that the stochastic transport equation admits a classical solution when the field is Lipschitz.

Chapter 3 is extracted from the paper *The Bismut-Elworthy-Li formula for mean-field stochastic differential equations*, [7]. As an application of the previous result and the fact that the solution of an SDE with Lipschitz drift is not only once but twice Malliavin differentiable and its Malliavin derivative is integrable in the sense of Skorokhod, we are able to derive a Bismut-Elworthy-Li formula for systems whose coefficients depend on the moments of the solution. This gives rise to expressions involving integration in the Skorokhod sense which usually requires higher order Malliavin regularity.

Chapter 4 is extracted from the paper *Optimal bounds for the densities of solutions of SDEs with measurable and path dependent drift coefficients*, [12] co-authored with P. Krühner. In this paper we study the problem of densities of Itô processes with additive noise, a classical problem which has plenty of applications, including finance, as for instance in the computations of *Greeks* using the so-called density method. As a start, in this chapter we develop a new method for finding sharp lower- and upper-bounds of densities of SDEs with irregular and path-dependent drift coefficients. We believe this method can be exploited to gain higher regularity of densities with very irregular drift coefficients.

Chapter 5 is extracted from the paper *Construction of Malliavin differentiable strong solutions of SDEs under an integrability condition on the drift without the Yamada-Watanabe principle*, [11] co-authored with S. Duedahl, T. Meyer-Brandis and F. Proske. It provides a new method for constructing Malliavin differentiable strong solutions of SDEs whose drift satisfies some integrability condition and in which no continuity or regularity is assumed. The method is a mixture of a compactness criterion of $L^2(\Omega)$ together with the classical idea of interchanging the irregular term with an expression involving the solution to the associated Kolmogorov's equation, which has better behaviour. As an application of the Malliavin differentiability of the solution we show a Bismut-Elworthy-Li formula.

Chapter 6 is extracted from the paper *Computing Deltas without derivatives*, [10] co-authored with S. Duedahl, T. Meyer-Brandis and F. Proske. In the same direction as previous chapter, but in the spirit of applying our research, we study the sensitivity of European, lookback and Asian options with respect to the initial condition, also known as *Delta*-sensitivity when both the drift coefficients of the underlying dynamics of the stock and the pay-off function are very irregular. Even in such a situation the price of the option is continuously differentiable and a classical *Delta* can be computed. We also analyse the numerical methods developed showing some simulations.

Chapter 7 is extracted from the paper *Strong existence and higher order differentiability of stochastic flows of fractional Brownian motion driven SDEs with singular drift*, [15] co-authored with T. Nilssen and F. Proske. Finally, we turn to the study of the problem of fractional Brownian noise. We also see this as the beginning of a study of other interesting problems involving fractional Brownian motion, as well as its implications in mathematical finance, as with the study of *Greeks* mentioned above. In this chapter we consider an SDE driven with fractional Brownian motion with very irregular drift coefficient and show that in this case one can also construct Malliavin differentiable strong solutions despite the fact that the solution process is neither Markovian nor a weak semimartingale, which means that classical techniques cannot be applied. To overcome these limitations we use the new method based on Malliavin calculus together with an *ad hoc* local-time variational calculus and an integration by parts formula. As an application, we also show that fractional Brownian motion regularises the flow associated to the SDE as the Hurst parameter gets smaller. In other words, one obtains a very regular flow with very irregular drift coefficient by using the adequate noise. We believe this phenomenon can be studied in other types of equations and stochastic partial differential equations.

Chapter 2

Malliavin and flow regularity of SDEs. Application to the study of densities and the stochastic transport equation

David R. Baños and Torstein Nilssen

Abstract: In this work we present a criterion for the regularity, in both space and Malliavin sense, of strong solutions to SDEs driven by Brownian motion. We conjecture that this criterion is optimal. As a consequence, we are able to improve the regularity of densities of such solutions.

We also apply these results to construct a classical solution to the stochastic transport equation when the drift is Lipschitz.

2.1 Introduction

This paper is mainly divided into two parts. First, we are interested in studying the regularity properties of the following Stochastic Differential Equation (SDE)

$$dX_t = b(t, X_t)dt + dB_t, \quad 0 \le t \le T, \quad X_0 = x \in \mathbb{R}^d, \tag{2.1}$$

where B_t, $t \in [0,T]$ is a d-dimensional Brownian motion and b is a measurable function such that a unique strong solution exists. Our goal is to analyse the regularity of strong solutions to (2.1) both in space and in the Malliavin sense. We give a criteria based on the regularity properties of b to obtain regularity properties of X_t, $t \in [0,T]$. Then we study the consequences of the aforementioned properties and we take two different directions. On one hand, the Malliavin regularity allows us to improve the regularity of densities of strong solutions. On the other hand, the regularity in space entitles us to study the associated Stochastic Transport Equation and gain more regularity on the solution. Namely, for b Lipschitz we are able to show that one obtains a classical solution to the Stochastic Transport Equation.

Considerable research in the direction of regularity of densities of solutions to SDEs has been done in the past years. There are well-known results on conditions for a density to be smooth when the coefficients are smooth, for example, in [90] or in the case of SDEs with boundary conditions in [65] under the so-called Hörmander's condition.

We highlight the work by S. Kusuoka and D. Stroock in [73] where the authors show that if $b \in C_b^{n+2}(\mathbb{R}^d)$, $n \geq 0$ then the density lies in $C_b^n(\mathbb{R}^d)$ using Sobolev inequalities associated to the H-derivative of the solution. Here, we improve the regularity of the density and skip the boundedness of b, instead we consider additive noise and provide an extension to a class of non-degenerate diffusion coefficients. In [72] S. Kusuoka also gives criteria for the law to be absolutely continuous with respect to the Lebesgue measure when drift coefficients are non-Lipschitz, his work is closely related to the findings of N. Bouleau and F. Hirsch in [24] where the authors show that for global Lipschitz coefficient a density exits, here we show that such density is Hölder continuous with exponent $\alpha < 1$. An improvent was given in dimension one in [51] where they show that for a Hölder continuous drift of, at most, linear growth and Hölder continuous diffusion coefficient the solution admits dentities at any given time.

Our technique is mainly based on Malliavin calculus and an a sharp estimate on the moments of the derivative of the flow associated to the solution, together with a strong result by V. Bally and L. Caramellino in [5] on the regularity of densities of random variables with sufficient Malliavin regularity. In addition, we also look at the regularity in space. As a consequence of the relationship between the Malliavin and Sobolev derivatives we are also able to give a criterion to determine the regularity of solutions to (2.1) in the Sobolev sense (locally) and show that such derivatives admit moments of any order. At the end of the section we also give an extension to more general diffusions.

The last application of the paper is devoted to the study of the Stochastic Transport Equation (STE) since it is closely related to the SDE (2.1) by the inverse of the flow of the solution. We use the results obtained in the first part of the paper to show that, for b Lipschitz, the solution is classical. Work in the direction of SPDE's and in particular the Stochastic Transport Equation has brought a lot of interests in the last years. In [47] F. Flandoli, M. Gubinelli and E. Priola study the well-posedness for Hölder-continuous drifts and show pathwise uniqueness of the weak solution. In [88], in dimension one, it is shown that when the drift is a step function then the solution to the transport equation is even once continuously differentiable.

2.2 Framework

In this section we recall some facts from Malliavin calculus and Sobolev spaces, which we aim at employing in Section 2.3 to analyse the regularity of densities of strong solutions of SDEs. See [90, 78, 79, 36] for a deeper insight on Malliavin Calculus. As for theory on Sobolev spaces the reader is referred to [75, 42].

2.2.1 Basic elements of Malliavin Calculus

In this Section we briefly elaborate a framework for Malliavin calculus.

Let $\{(\Omega, \mathcal{F}, P)\,; H\}$ be a Gaussian probability space, that is $(\Omega, \mathcal{F}, P)$ is a probability space and H a separable closed subspace of centered Gaussian random variables of $L^2(\Omega)$, which generate the σ-field $\mathcal{F}$. Denote by D the derivative operator acting on elementary smooth random variables in the sense that

$$D(f(h_1,\dots,h_n)) = \sum_{i=1}^{n} \partial_i f(h_1,\dots,h_n)h_i, \quad h_i \in H, \quad f \in C_b^\infty(\mathbb{R}^n).$$

Further let $\mathbb{D}^{k,p}(\Omega)$, $k,p \geq 1$ be the closure of the family of elementary smooth random variables with respect to the norm

$$\|F\|_{\mathbb{D}^{k,p}(\Omega)} := \|F\|_{L^p(\Omega)} + \sum_{i=1}^{k} \left\| D\overset{i)}{\cdots}DF \right\|_{L^p(\Omega; H\otimes\overset{i)}{\cdots}\otimes H)}.$$

Our framework will rely on the special case where $H = L^2([0,T];\mathbb{R}^d)$, then we have that the Malliavin derivative is a process $\{D_tF\}_{t\in[0,T]}$ in $L^2(\Omega \times [0,T];\mathbb{R}^d)$ defined as

$$D_tF = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f\left(\int_0^T h_1(u)dW_u, \dots, \int_0^T h_n(u)dW_u\right) h_i(t)$$

and in this case we take the closure w.r.t. the norm

$$\|F\|_{\mathbb{D}^{k,p}(\Omega)} = E[|F|^p]^{1/p} + \sum_{i=1}^{k} E\left[\int_0^T \cdots \int_0^T \|D_{t_1}\cdots D_{t_i}F\|^p dt_1 \cdots t_i\right]^{1/p}$$

where $\|\cdot\|$ denotes any norm in $\mathbb{R}^{d\times\overset{i)}{\cdots}\times d}$.

The operator $D\overset{k)}{\cdots}D$ is then a closed operator from $\mathbb{D}^{k,p}(\Omega)$ to $L^p(\Omega\times[0,T]^k;\mathbb{R}^{d\times\overset{k)}{\cdots}\times d})$ for all $p\geq 1$. Moreover, for $p \leq q$ and $k \leq l$ we have

$$\|F\|_{\mathbb{D}^{k,p}(\Omega)} \leq \|F\|_{\mathbb{D}^{l,q}(\Omega)}$$

and as a consequence

$$\mathbb{D}^{k+1,p}(\Omega) \hookrightarrow \mathbb{D}^{k,q}(\Omega)$$

if $k \geq 0$ and $p > q$.

We shall say that a random variable is k-times *Malliavin differentiable* with derivatives in $L^p(\Omega)$, $p \geq 1$ if it lies on $\mathbb{D}^{k,p}(\Omega)$.

Finally, we have the chain-rule for the Malliavin derivative. Let $\varphi : \mathbb{R}^m \to \mathbb{R}^m$ be a function such that

$$|\varphi(x) - \varphi(y)| \leq K|x-y|$$

for any $x, y \in \mathbb{R}^m$. Suppose $F = (F^1,\dots,F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,2}(\Omega)$. Then $\varphi(F) \in \mathbb{D}^{1,2}(\Omega)$ and there exists a random vector $G = (G_1,\dots,G_m)$ bounded by K such that

$$D\varphi(F) = \sum_{i=1}^{m} G_i DF^i.$$

In particular if φ' exists and the law of the reandom variable F is absolutely continuous with respect to the Lebesgue measure, then $G = \varphi'(F)$.

2.2.2 Basic facts of theory on Sobolev spaces

In this section we concisely review some basic facts about theory on Sobolev spaces.

Let U be an open bounded subset of $\mathbb{R}^d$. Fix $1 \leq p \leq \infty$ and let $k \geq 0$ an integer. The Sobolev space $W^{k,p}(U)$ is composed by all locally L^p-integrable functions $u : U \to \mathbb{R}^d$ such that for any multiindex α with $|\alpha| \leq k$, then $D^\alpha u$ exists in the weak sense and belongs to $L^p(U)$.

We endow the space $W^{k,p}(U)$ with the topology generated by the norm

$$\|u\|_{W^{k,p}(U)} := \left(\sum_{|\alpha| \leq k} \int_U |D^\alpha u|^p dx \right)^{1/p}, \quad 1 \leq p < \infty$$

or

$$\|u\|_{W^{k,\infty}(U)} := \sum_{|\alpha| \leq k} \operatorname*{ess\,sup}_U |D^\alpha u|, \quad p = \infty.$$

The following relations will be of high relevance for our purposes. For $1 \leq p < q \leq \infty$, $k > l$ such that $(k-l)p < d$ and

$$\frac{1}{q} = \frac{1}{p} - \frac{k-l}{d}$$

then we have the following continuous embedding

$$W^{k,p}(\mathbb{R}^d) \hookrightarrow W^{l,q}(\mathbb{R}^d). \tag{2.2}$$

Also, we have the following embedding as a consequence of Morrey's inequality; if $\frac{k-r-\alpha}{d} = \frac{1}{p}$ with $\alpha \in (0,1)$ then

$$W^{k,p}(\mathbb{R}^d) \hookrightarrow C^{r,\alpha}(\mathbb{R}^d). \tag{2.3}$$

Essentially, this means that if we have enough Sobolev regularity then we may expect some continuous classical derivatives up to some order.

We will though use $\frac{\partial}{\partial x}$ to denote differentiation in both the weak and classical sense when the context is clear.

2.2.3 Shuffles

Let $m, n \in \mathbb{N}_0$ and denote by $S_m = \{\sigma : \{1, \dots, m\} \to \{1, \dots, m\}\}$ the set of permutations of length m. Define the set of *shuffle permutations* of length $m+n$ as

$$S(m,n) := \{\sigma \in S_{m+n} : \sigma(1) < \cdots < \sigma(m), \sigma(m+1) < \cdots < \sigma(m+n)\}.$$

Fix $s, t \in [0,T]$ with $s < t$ and define the m-dimensional subset of $[0,T]^m$

$$\Lambda^m_{s,t} := \{(u_1, \dots, u_m) \in [0,T]^m : s < u_1 < \cdots < u_m < t\}.$$

Let $f_i : [0,T] \to \mathbb{R}$, $i = 1, \ldots, m+n$ be integrable functions. Then, we have

$$\begin{aligned}\left(\int_{\Lambda^m_{s,t}} \prod_{i=1}^{m} f_i(u_i) du_1 \cdots du_m\right)\left(\int_{\Lambda^n_{s,t}} \prod_{i=m+1}^{m+n} f_i(u_i) du_{m+1} \cdots du_{m+n}\right)& \\ = \sum_{\sigma^{-1}\in S(m,n)} \int_{\Lambda^{m+n}_{s,t}} \prod_{i=1}^{m+n} f_{\sigma(i)}(u_i) du_1 \cdots du_{m+n}&\end{aligned} \tag{2.4}$$

since

$$\begin{aligned}&\left\{s < u_1 < \cdots < u_m < t,\, s < u_{m+1} < \cdots < u_{m+n} < t\right\} \\ &\quad = \bigcup_{\sigma\in S(m,n)}\!\!\!\!\!\!\cdot\;\; \left\{(w_1, \ldots, w_{m+n}) \in [0,T]^{m+n} :\, s < w_{\sigma(1)} < \cdots < w_{\sigma(m+n)} < t\right\},\end{aligned}$$

which can also be found in [77, Theorem 2.15].

We will also need the following formula. Given an index $r \in \mathbb{N}$ such that $1 \le r \le n$. Introduce the subset $S_r(m,n)$ of $S(m,n)$ defined as

$$\begin{aligned}S_r(m,n) := \Big\{&\sigma \in S(m,n) :\, \sigma(1) < \cdots < \sigma(m),\, \sigma(m+1) < \cdots < \sigma(m+r-1), \\ &\sigma(l) = l,\, m+r \le l \le m+n\Big\}.\end{aligned}$$

We have

$$\begin{aligned}\int_{\Lambda^n_{s,t}\times\Lambda^m_{s,u_{m+r}}} \prod_{i=1}^{m+n} f_i(u_i) du_1 \cdots du_{m+n} &= \int_{\substack{s<u_1<\cdots<u_m<u_{m+r} \\ s<u_{m+1}<\cdots<s_{m+n}<t}} \prod_{i=1}^{m+n} f_i(u_i) du_1 \cdots du_{m+n} \\ &= \sum_{\sigma^{-1}\in S_r(m,n)} \int_{s<w_1<\cdots<w_{m+n}<t} \prod_{i=1}^{m+n} f_i(w_i) dw_1 \cdots dw_{m+n}.\end{aligned} \tag{2.5}$$

Observe also that

$$\#S(m,n) = \frac{(m+n)!}{m!n!}$$

where $\#$ denotes the number of elements in the given set. Then by using Stirling's approximation, one can show that

$$\#S(m,n) \le C^{m+n}$$

for a large enough constant $C > 0$. Moreover,

$$\#S_r(m,n) \le \#S(m,n).$$

2.3 Malliavin and flow regularity of strong solutions of SDEs

Consider the *stochastic differential equation* (SDE) given by

$$\begin{cases} dX_t = b(t, X_t)dt + dB_t, \\ X_0 = x \in \mathbb{R}^d, \end{cases} \tag{2.6}$$

where the drift coefficient $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable function and B_t is a d-dimensional Brownian motion defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P)$ where the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the one generated by B_t, $t \in [0,T]$ augmented by the P-null sets.

If b is of linear growth and Lipschitz continuous it is well-known that there exists a unique global strong solution to the SDE (2.6) which belongs to $\mathbb{D}^{1,2}(\Omega)$. In fact, under more relaxed conditions on b one has the same result, see for instance [81], [88].

In this section we are concerned with the regularity of the solution in the Malliavin sense in terms of the regularity of b. We will assume the following hypotheses for b, for every $(t,x) \in [0,T] \times \mathbb{R}^d$

$$\begin{gathered} |b(t,x)| \leq C(1+|x|), C > 0, \\ Db(t,\cdot), D^2b(t,\cdot), \dots, D^kb(t,\cdot) \in L^\infty(\mathbb{R}^d) \end{gathered} \tag{H}$$

for some $k \geq 1$ where here, the derivatives are understood in the weak sense. In particular, b is $k-1$ times continuously differentiable in virtue of the Sobolev embedding (2.3) and equation (2.6) admits a unique strong solution.

Before we proceed to the main statements of this section we need two preliminary results which are essential for our targets.

Lemma 2.1. *Let $\{b_n\}_{n\geq 0}$ be a sequence of compactly supported smooth functions approximating b a.e. in $t \in [0,T]$ and $x \in \mathbb{R}^d$ such that $\sup_{n\geq 0} |b_n(t,x)| \leq C(1+|x|)$, all $x \in \mathbb{R}^d$ and $t \in [0,T]$. Then for any compact subset $K \subset \mathbb{R}^d$ there exists an $\varepsilon > 0$ such that*

$$\sup_{x\in K}\sup_{n\geq 0} E\left[\mathcal{E}\left(\int_0^T b_n(u, B_u^x)dB_u\right)^{1+\varepsilon}\right] < \infty. \tag{2.7}$$

where $B_t^x := x + B_t$ and $\mathcal{E}(M_t)$ denotes the Doléans-Dade exponential of a martingale M_t. In particular we also have

$$\sup_{x\in K} E\left[\mathcal{E}\left(\int_0^T b(u, B_u^x)dB_u\right)^{1+\varepsilon}\right] < \infty. \tag{2.8}$$

Proof. Indeed, write

$$
\begin{aligned}
&E\left[\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)^{1+\varepsilon}\right]=\\
&=E\left[\exp\left\{\int_0^T(1+\varepsilon)b_n(u,B_u^x)dB_u-\frac{1}{2}\int_0^T(1+\varepsilon)|b_n(u,B_u^x)|^2du\right\}\right]\\
&=E\Bigg[\exp\Bigg\{\int_0^T(1+\varepsilon)b_n(u,B_u^x)dB_u-\frac{1}{2}\int_0^T(1+\varepsilon)^2|b_n(u,B_u^x)|^2du\\
&+\frac{1}{2}\int_0^T\varepsilon(1+\varepsilon)|b_n(u,B_u^x)|^2du\Bigg\}\Bigg]\\
&=E\left[\exp\left\{\frac{1}{2}\int_0^T\varepsilon(1+\varepsilon)|b_n(u,X_u^{\varepsilon,x})|^2du\right\}\right]
\end{aligned}
$$

where the last step follows from Girsanov's theorem and here $X_t^{\varepsilon,x}$ is a solution of the following SDE

$$
\begin{cases}
dX_t^{\varepsilon,x}=(1+\varepsilon)b_n(t,X_t^{\varepsilon,x})dt+dB_t, & t\in[0,T]\\
X_0^{\varepsilon,x}=x.
\end{cases}
$$

Observe that, since b has at most linear growth, we have

$$
|X_t^{\varepsilon,x}|\le|x|+C(1+\varepsilon)\int_0^t(1+|X_u^{\varepsilon,x}|)du+|B_t|
$$

for every $t\in[0,T]$. Then Grönwall's inequality gives

$$
|X_t^{\varepsilon,x}|\le(|x|+C(1+\varepsilon)T+|B_t|)\,e^{C(1+\varepsilon)T}, \tag{2.9}
$$

and the sublinearity of b_n and the estimate (2.9) give

$$
|b_n(u,X_u^{\varepsilon,x})|\le C_{\varepsilon,T}\,(1+|x|+|B_t|)
$$

where $C_{\varepsilon,T}$ denotes the collection of all constants depending on ε,T.

As a result,

$$
\begin{aligned}
E\Big[\exp\Big\{\varepsilon(1+\varepsilon)\int_0^T|b_n(u,X_u^{\varepsilon,x})|^2du\Big\}\Big]&\le E\left[\exp\left\{\tilde{C}_{\varepsilon,T}\int_0^T(1+|x|+|B_u|)^2\,du\right\}\right]\\
&\le e^{\tilde{C}_{\varepsilon,T}(1+|x|)^2}E\left[\exp\left\{\tilde{C}_{\varepsilon,T}(1+|x|)\int_0^T(|B_u|+|B_u|^2)du\right\}\right]
\end{aligned}
$$

where $\tilde{C}_{\varepsilon,T}>0$ is a constant such that $\lim_{\varepsilon\searrow 0}\tilde{C}_{\varepsilon,T}=0$. Clearly, for every compact set $K\subset\mathbb{R}^d$ we can choose $\varepsilon>0$ small enough such that

$$
\sup_{x\in K}\sup_{n\ge0}E\Big[\exp\Big\{\varepsilon(1+\varepsilon)\int_0^T|b_n(u,X_u^{\varepsilon,x})|^2du\Big\}\Big]<\infty.
$$

□

Remark 2.2. *We point out that the finite dimensional laws of the strong solution of (2.6) are absolutely continuous with respect to the Lebesgue measure. To see this, let A denote a set with null Lebesgue measure. Then since b is of, at most, linear growth, by Girsanov's theorem, see e.g. [62, Proposition 5.3.6] and Lemma 2.8 one has for some $\varepsilon > 0$ small enough*

$$\begin{aligned}P(X_t \in A) &\leq E\left[\mathbf{1}_{\{B_t^x \in A\}}\mathcal{E}\left(\int_0^T b(u, B_u^x)dB_u\right)\right] \\ &\leq C_\varepsilon P(B_t^x \in A)^{\frac{\varepsilon}{1+\varepsilon}} \\ &= 0.\end{aligned}$$

Next, we give a crucial estimate for the proof of our main results.

Proposition 2.3. *Let B be a d-dimensional Brownian Motion starting from $z_0 \in \mathbb{R}^d$ and $b_1, \ldots, b_m$ be compactly supported continuously differentiable functions $b_i : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ for $i = 1, 2, \ldots m$. Let $\alpha_i \in \{0,1\}^d$ be a multiindex such that $|\alpha_i| \leq 1$ for $i = 1, 2, \ldots, m$. Then there exists a universal constant C (independent of $\{b_i\}_i$, m, and $\{\alpha_i\}_i$) such that*

$$\left|E\left[\int_{t_0<t_1<\cdots<t_m<t}\left(\prod_{i=1}^m D^{\alpha_i} b_i(t_i, B_{t_i})\right)dt_1 \ldots dt_m\right]\right| \leq \frac{C^m \prod_{i=1}^m \|b_i\|_\infty (t-t_0)^{m/2}}{\Gamma(\frac{m}{2}+1)} \tag{2.10}$$

for every $t_0, t \in [0,T]$ where Γ is the Gamma-function. Here D^{α_i} denotes the partial derivative with respect to the j'th space variable, where j is the position of the 1 in α_i.

Proof. Observe that $|\alpha_i| \leq 1$, that is we allow for the possibility of some of the functions in (2.10) not being differentiated. For the case when $|\alpha_i| = 1$ for all $i = 1, \ldots, m$ a detailed proof of the estimate can be already found in [81, Proposition 3.7]. Here, we will show that the result still holds when we have less derivatives involved in the integrand by use of the same ideas as in [81], which on the other side seems intuitive.

So, without loss of generality, assume that $\|b_i\|_\infty \leq 1$ for $i = 1, 2 \ldots, m$. Denote by

$$\Lambda_m(t_i, t) := \{(t_{i+1}, \ldots, t_m) \in [0,T]^m,\, t_i < t_{i+1} < \ldots, t_m < t\}$$

for any $i = 0, \ldots, m-1$. Moreover, denote by $z = (z^{(1)}, \ldots z^{(d)})$ a generic element of $\mathbb{R}^d$ and by $|\cdot|$ the usual Euclidean norm. With $P(t,z) = (2\pi t)^{-d/2}e^{-|z|^2/2t}$ and the independence of the increments of the Brownian motion write the left hand side in (2.10) as

$$\left|\int_{\Lambda_m(t_0,t)}\int_{\mathbb{R}^{dm}}\prod_{i=1}^m g_i(t_i, z_i)P(t_i - t_{i-1}, z_i - z_{i-1})dz_1 \ldots dz_m dt_1 \ldots dt_m\right| \tag{2.11}$$

where

$$g_i \in \{b_i, D^{\alpha_i} b_i\},\ i = 1, \ldots, m \tag{2.12}$$

In order to make notation more tangible, we assume, without loss of generality, that the integrand of (2.11) is of the form

$$\prod_{i=1}^{m} g_i = G_{1,j_1} G_{j_1+1,j_2} \quad \cdots \quad G_{j_{k-1}+1,j_k}$$

with indexes $j_1, \ldots, j_k$ where $j_k = m$ and each $G_{j_{i-1}+1,j_i}$ for each $i = 1, \ldots, k$ represents the (simplest) block of size $j_i - j_{i-1}$, $(j_0 = 0)$ where all functions are equal in the sense described in (2.12) and k is the total number of blocks. In other words, for each $i = 1, \ldots, k$

$$G_{j_{i-1}+1,j_i}(t_{j_{i-1}+1}, \ldots, t_{j_i}, z_{j_{i-1}+1}, \ldots, z_{j_i}) = \prod_{l=j_{i-1}+1}^{j_i} g_l(t_l, z_l)$$

with either all $g_l = b_l$ or $g_l = D^{\alpha_l} b_l$, $l = j_{i-1}+1, \ldots, j_i$ and $i = 1, \ldots, k$. Assume, for instance, that k is even and that the blocks $G_{j_{i-1}+1,j_i}$ with odd $i = 1, 3, \ldots, k-1$ consist of $b_{j_{i-1}+1}, \ldots, b_{j_i}$ and the ones with even $i = 2, 4, \ldots, k$ consist of $D^{\alpha_{j_{i-1}+1}} b_{j_{i-1}+1}, \ldots, D^{\alpha_{j_i}} b_{j_i}$. Introduce the notation

$$\begin{aligned} &J_m^\alpha(t_0, t, z_0) \\ &\quad = \int_{\Lambda_m(t_0,t)} \int_{\mathbb{R}^{dm}} \prod_{i=1}^{k} G_{j_{i-1}-1,j_i}(t_{j_{i-1}-1,\ldots,t_{j_i}}, z_{j_{i-1}-1}, \ldots, z_{j_i}) \prod_{j=1}^{m} P(t_j - t_{j-1}, z_j - z_{j-1}) dz dt \end{aligned}$$

where $z = (z_1, \ldots, z_m) \in \mathbb{R}^{dm}$, $t = (t_1, \ldots, t_m) \in \mathbb{R}^m$ and $\alpha = (\alpha_{j_1}, \ldots, \alpha_{j_k}) \in \{0,1\}^{md}$ such that $\alpha_{j_i} := (\alpha_{j_{i-1}+1}, \ldots, \alpha_{j_i}) = \vec{0}$ for i odd. We can then write J_m^α as follows

$$\begin{aligned} &J_m^\alpha(t_0, t, z_0) \\ &= \int_{\Lambda_m(t_0,t)} \int_{\mathbb{R}^{dm}} \prod_{\substack{i=1 \\ i \text{ odd}}}^{k} \prod_{l=j_{i-1}+1}^{j_i} b_l(t_l, z_l) \prod_{\substack{i=1 \\ i \text{ even}}}^{k} \prod_{l=j_{i-1}+1}^{j_i} D^{\alpha_l} b_l(t_l, z_l) \prod_{i=1}^{m} P(t_i - t_{i-1}, z_i - z_{i-1}) dz dt. \end{aligned}$$

We shall show that $|J_m^\alpha(t_0, t, 0)| \leq C^m (t - t_0)^{m/2} / \Gamma(m/2 + 1)$, thus proving the proposition. To do this, we will use integration by parts to shift the derivatives onto the Gaussian kernels. We are only interested in transferring the D^{α_l} from the blocks $G_{j_{i-1}+1,j_i}$ corresponding to the even $i = 1, \ldots, k$. Observe that, given an even i, the derivatives involved in the block $G_{j_{i-1}+1,j_i}$ influence the block of Gaussian kernels $\prod_{l=j_{i-1}+1}^{j_i} P(t_l - t_{l-1}, z_l - z_{l-1})$ and the first kernel from the next block, i.e. $P(t_{j_i+1} - t_{j_i}, z_{j_i+1} - z_{j_i})$. Thus, we see that we can write $J_m^\alpha(t_0, t, z_0)$ as

$$\begin{aligned} J_m^\alpha(t_0, t, z_0) = &(-1)^{|\alpha|} \int_{\Lambda_m(t_0,t)} \int_{\mathbb{R}^{dm}} \prod_{i=1}^{m} b_i(t_i, z_i) D^{\alpha_{j_1+1}} \cdots D^{\alpha_{j_2}} D^{\alpha_{j_3+1}} \cdots D^{\alpha_{j_4}} \cdots \\ &\cdots D^{\alpha_{j_{k-1}+1}} \cdots D^{\alpha_{j_k}} \big[P(t_1 - t_0, z_1 - z_0) \cdots P(t_m - t_{m-1}, z_m - z_{m-1}) \big] dz dt. \end{aligned} \tag{2.13}$$

We will proceed by introducing an alphabet, as the authors did in [81, Proposition 3.7], as follows: consider $\mathcal{A}(\alpha_1, \ldots, \alpha_m) = \{P, D^{\alpha_1} P, \ldots, D^{\alpha_m} P, D^{\alpha_1} D^{\alpha_2} P, \ldots D^{\alpha_{m-1}} D^{\alpha_m} P\}$ where

D^{α_i}, $D^{\alpha_i}D^{\alpha_{i+1}}$ denote the derivatives in z on $P(t,z)$. We will only need a special type of strings, and we say that a string is *allowed* if, when all the $D^{\alpha_i}P$'s are removed from the string, a string of the form $P \cdot D^{\alpha_s}D^{\alpha_{s+1}}P \cdot P \cdot D^{\alpha_{s+1}}D^{\alpha_{s+2}}P \cdots P \cdot D^{\alpha_r}D^{\alpha_{r+1}}P$ for $s \geq 1, r \leq m-1$ remains. Also, we will require that the first derivatives $D^{\alpha_i}P$ are written in an increasing order with respect to i.

We see that the derivatives in the integrand in (2.13) produce sums of differentiated Gaussian kernels (because of Leibniz product rule) and one may observe that the resulting kernels in each of these summands are of the form

$$P \overset{j_1)}{\cdots} P\, S_{j_1+1,j_2+1}\, P \overset{j_3-j_2-1)}{\cdots} P\, S_{j_3+1,j_4+1}\, P \overset{j_5-j_4-1)}{\cdots} P\, \cdots\, S_{j_{k-1}+1,m} \tag{2.14}$$

where each $S_{j_{i-1}+1,j_i+1} := S_{j_{i-1}+1} \cdots S_{j_i} S_{j_i+1}$, for even i, is an allowed string in $\mathcal{A}(\alpha_{j_{i-1}+1}, \ldots, \alpha_{j_i}, \alpha_{j_i+1})$ and for the last one we agree that $S_{j_k+1} = 1$. Denote by $S \in \mathcal{B}(\alpha)$ strings of the form described in (2.14), in other words, strings such that when removing all $P \cdots P$'s from the string (2.14), we remain with an allowed string.

Define for a string $S = S_1 \cdots S_m \in \mathcal{B}(\alpha)$

$$I_S^\alpha(t_0,t,z_0) = \int_{\Lambda_m(t_0,t)} \int_{\mathbb{R}^{dm}} \prod_{i=1}^m b_i(t_i,z_i) S_i(t_i - t_{i-1}, z_i - z_{i-1}) dz dt.$$

This allows us to write

$$J_m^\alpha(t_0,t,z_0) = \sum_{l=1}^{2^{N-1}} \epsilon_l I_{S_l}^{\alpha_l}(t_0,t,z_0)$$

where each ϵ_l is either -1 or 1, each S_l is an string in $\mathcal{B}(\alpha)$ and

$$N = \sum_{\substack{i=1,\ldots,k \\ i \text{ even}}} j_i - j_{i-1}.$$

Observe that 2^{N-1} is the total number of summands so if we only have one block ($k=1$) then $j_1 = m$ and $N = 2^{m-1}$. The proof can be easily reduced to the case in [81, Lemma 3.8].

We know that the estimate holds only for allowed strings $S \in \mathcal{A}(\alpha)$ due to [81, Proposition 3.7]. So for even $1 \leq i \leq k$ we have for some constant $C > 0$

$$\begin{aligned} &\int_{\Lambda_m(t_{j_{i-1}},t)} \int_{\mathbb{R}^{(j_i - j_{i-1})d}} S_{j_{i-1}+1}(t_{j_{i-1}+1} - t_{j_{i-1}}, z_{j_{i-1}+1} - z_{j_{i-1}}) \cdots S_{j_i}(t_{j_i} - t_{j_i-1}, z_{j_i} - z_{j_i-1}) \\ &\times S_{j_i+1}(t_{j_i+1} - t_{j_i}, z_{j_i+1} - z_{j_i}) dz_{j_i+1} \cdots dz_{j_{i-1}+1} dt_{j_i+1} \cdots dt_{j_{i-1}+1} \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \leq C^{j_i - j_{i-1}} \frac{|t_{j_i} - t_{j_{i-1}}|^{\frac{j_i - j_{i-1}}{2}}}{\Gamma\left(\frac{j_i - j_{i-1}}{2} + 1\right)}. \end{aligned}$$

This and the fact that $\int_{\mathbb{R}^d} P(t,z)dz = 1$ give

$$|I_S^\alpha(t_0,t,z_0)| \leq \prod_{\substack{i=1,\ldots,k \\ i \text{ even}}} \frac{C^{j_i - j_{i-1}}}{\Gamma\left(\frac{j_i - j_{i-1}}{2} + 1\right)} \int_{\Lambda_m(t_0,t)} \prod_{\substack{i=1,\ldots,k \\ i \text{ even}}} |t_{j_i} - t_{j_{i-1}}|^{\frac{j_i - j_{i-1}}{2}} dt_m \cdots dt_1.$$

Further,

$$\int_{\Lambda_m(t_0,t)} \prod_{\substack{i=1,\dots,k\\ i \text{ even}}} |t_{j_i} - t_{j_{i-1}}|^{\frac{j_i - j_{i-1}}{2}} dt_m \cdots dt_1 \leq |t-t_0|^{\frac{1}{2}\sum_{i \text{ even}} j_i - j_{i-1}} \frac{|t-t_0|^m}{\Gamma(m+1)}$$
$$\leq M^m \frac{|t-t_0|^{m/2}}{\Gamma(m+1)}$$

for some constant M.

Finally, since

$$\Gamma\left(\frac{m}{2}+1\right) \leq \prod_{\substack{i=1,\dots,k\\ i \text{ even}}} \Gamma\left(\frac{j_i - j_{i-1}}{2}+1\right) \Gamma\left(m+1\right)$$

the result follows. □

We turn now to one of the main results of this section.

Theorem 2.4. *Let X_t, $t \in [0,T]$ denote the solution to equation (2.6) with $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ a function satisfying hypotheses (H) with $k = 1$, i.e. of linear growth with bounded weak derivative, then we have $X_t \in \mathbb{D}^{2,p}(\Omega)$ for all $p \geq 1$. In particular the result holds if b is (globally) Lipschitz.*

Proof. In order to carry out the proof of Theorem 2.4, we use the following result in [90, Proposition 1.5.5.].

Proposition 2.5. *Let $\{X_n\}_{n\geq 0}$ a sequence of random variables such that $X_n \to X$ in $L^p(\Omega)$, $p \geq 1$ and such that for $k \geq 1$*

$$\sup_{n\geq 0} \|X_n\|_{\mathbb{D}^{k,p}(\Omega)} < \infty,$$

then $X \in \mathbb{D}^{k,p}(\Omega)$.

We start with the proof of Theorem 2.4 by showing that the solution X_t of (2.6) can be approximated by random variables in $L^p(\Omega)$ for every $t \in [0,T]$.

We have $b' \in L^\infty(\mathbb{R}^d)$ and b has linear growth, i.e. there is $C > 0$ such that

$$|b(t,x)| \leq C(1+|x|)$$

for every $t \in [0,T]$ and $x \in \mathbb{R}^d$. Then we can approximate b a.e. in $t \in [0,T]$ and $x \in \mathbb{R}^d$ by a sequence of functions $\{b_n\}_{n\geq 1} \subset C^2(\mathbb{R}^d)$ such that $\sup_{n\geq 0} |b_n(t,x)| \leq C(1+|x|)$ and $\sup_n \|b_n'\|_\infty < \infty$. For each $t \in [0,T]$, denote by X_t^n the sequence of random variables in $L^p(\Omega)$ solution to equation (2.6) with drift coefficient b_n. Then

$$X_t^n = x + \int_0^t b_n(u, X_u^n) du + B_t.$$

Denote by p_{X_t} the density of X_t for a fixed $t \in [0,T]$ from Remark 2.2. Denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^d$, then

$$
\begin{aligned}
E\big[|X_t^n - X_t|^p\big] &= E\left[\left|\int_0^t (b_n(u,X_u^n) - b(u,X_u))\,du\right|^p\right] \\
&\leq (2t)^{p-1} E\left[\int_0^t |b_n(u,X_u) - b(u,X_u)|^p du\right] + (2t)^{p-1} E\left[\int_0^t |b_n(u,X_u^n) - b_n(u,X_u)|^p du\right] \\
&\leq (2t)^{p-1} E\left[\int_0^t |b_n(u,X_u) - b(u,X_u)|^p du\right] + (2t)^{p-1} \|b_n'\|_\infty^p E\left[\int_0^t |X_u^n - X_u|^p du\right].
\end{aligned}
$$

Using Grönwall's inequality we obtain

$$
\begin{aligned}
E\left[|X_t^n - X_t|^p\right] &\leq (2t)^{p-1} \exp\left\{(2t)^{p-1} t \sup_k \|b_k'\|_\infty^p\right\} E\left[\int_0^t |b_n(u,X_u) - b(u,X_u)|^p du\right] \\
&\leq C \int_0^t \int_{\mathbb{R}^d} |b_n(u,x) - b(u,x)|^p p_{X_u}(x) dx du
\end{aligned}
$$

for a constant $C > 0$ independent of n. Then Lebesgue's dominated convergence theorem gives the $L^p(\Omega)$-convergence.

Let us now proceed with the proof that the random variables X_t^n are bounded in $\mathbb{D}^{2,p}(\Omega)$ for every $p \geq 1$: Fix $s_1, t \in [0,T]$, $s_1 \leq t$. Then

$$
D_{s_1} X_t^n = \mathcal{I}_d + \int_{s_1}^t b_n'(u, X_u^n) D_{s_1} X_u^n du. \tag{2.15}
$$

The above equations for $D_{s_1} X_u^n$, $n \geq 1$, are linear equations with matrix-valued unknowns. Since each b_n is smooth we have a unique solution of (2.15). Again, for notational convenience we denote by $\Lambda_m(s,t) := \{(u_1, \dots, u_m) \in [0,T]^m : \ s < u_1 < \cdots < u_m < t\}$ the m-dimensional simplex. Then using a Picard iteration argument we may write the solution of (2.15) as a series expansion as follows

$$
D_{s_1} X_t^n = \mathcal{I}_d + \sum_{m \geq 1} \int_{\Lambda_m(s_1,t)} b_n'(u_1, X_{u_1}^n) \cdots b_n'(u_m, X_{u_m}^n) du_1 \cdots du_m. \tag{2.16}
$$

To see that the above expression is indeed the solution of (2.15) just make the following observation

$$
\frac{d}{dt} D_{s_1} X_t^n = b_n'(t, X_t^n) \left(\mathcal{I}_d + \sum_{m \geq 1} \int_{\Lambda_m(s_1,t)} b_n'(u_1, X_{u_1}^n) \cdots b_n'(u_m, X_{u_m}^n) du_1 \cdots du_m\right).
$$

Take now $0 \leq s_1 \leq s_2 \leq t$. Then

$$
D_{s_2} D_{s_1} X_t^n = \sum_{m \geq 1} \int_{\Lambda_m(s_2,t)} D_{s_2}\left[b_n'(u_1, X_{u_1}^n) \cdots b_n'(u_m, X_{u_m}^n)\right] du_1 \cdots du_m. \tag{2.17}
$$

We expand the integrand of (2.17) using Leibniz's rule as follows

$$D_{s_2}\left[b'_n(u_1,X^n_{u_1})\cdots b'_n(u_m,X^n_{u_m})\right]=\sum_{r=1}^m b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})D_{s_2}X^n_{u_r}\cdots b'_n(u_m,X^n_{u_m}).$$

Remark 2.6. *We recall here that $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ so $Db(t,\cdot):\mathbb{R}^d\to L(\mathbb{R}^d,\mathbb{R}^d)$ and $Db(t,x)\in L(\mathbb{R}^d,\mathbb{R}^d)$. The second derivative is then $D^2b(t,\cdot):\mathbb{R}^d\to L(\mathbb{R}^d,L(\mathbb{R}^d,\mathbb{R}^d))$ so $D^2b(t,x):L(\mathbb{R}^d,L(\mathbb{R}^d,\mathbb{R}^d))\cong L^2(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ denoting by $L^2(\mathbb{R}^d\times\mathbb{R}^d,\mathbb{R}^d)$ the bilinear forms from $\mathbb{R}^d\times\mathbb{R}^d$ into $\mathbb{R}^d$.*

Inserting the representation (2.16) for $D_{s_2}X^n_{u_r}$ in this case we have that the above quantity can be written as

$$\sum_{r=1}^m b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})\left(\mathcal{I}_d+\sum_{m\geq 1}\int_{\Lambda_m(s_2,u_r)} b'_n(v_1,X^n_{v_1})\cdots b'_n(v_m,X^n_{v_m})dv_1\cdots dv_m\right)$$
$$\times b'_n(u_{r+1},X^n_{u_{r+1}})\cdots b'_n(u_m,X^n_{u_m}).$$

Altogether

$$\begin{aligned}
D_{s_2}D_{s_1}X^n_t &= \sum_{m_1\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\sum_{r=1}^{m_1} b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})\\
&\times\left(\mathcal{I}_d+\sum_{m_2\geq 1}\int_{\Lambda_{m_2}(s_2,u_r)} b'_n(v_1,X^n_{v_1})\cdots b'_n(v_{m_2},X^n_{v_{m_2}})dv_1\cdots dv_{m_2}\right)\\
&\times b'_n(u_{r+1},X^n_{u_{r+1}})\cdots b'_n(u_{m_1},X^n_{u_{m_1}})du_1\cdots du_{m_1}.\\
&= \sum_{m_1\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\sum_{r=1}^{m_1} b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})\cdots b'_n(u_{m_1},X^n_{u_{m_1}})du_1\cdots du_{m_1}\\
&+\sum_{m_1\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\sum_{r=1}^{m_1} b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})\\
&\times\left(\sum_{m_2\geq 1}\int_{\Lambda_{m_2}(s_2,u_r)} b'_n(v_1,X^n_{v_1})\cdots b'_n(v_{m_2},X^n_{v_{m_2}})dv_1\cdots dv_{m_2}\right)\\
&\times b'_n(u_{r+1},X^n_{u_{r+1}})\cdots b'_n(u_{m_1},X^n_{u_{m_1}})du_1\cdots du_{m_1}.
\end{aligned}$$

We reallocate terms by dominated convergence and respecting the order of matrices

$$D_{s_2}D_{s_1}X^n_t=\sum_{m_1\geq 1}\sum_{r=1}^{m_1}\int_{\Lambda_{m_1}(s_2,t)} b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})\cdots b'_n(u_{m_1},X^n_{u_{m_1}})du_1\cdots du_{m_1} \tag{2.18}$$
$$+\sum_{m_1\geq 1}\sum_{r=1}^{m_1}\sum_{m_2\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\int_{\Lambda_{m_2}(s_2,u_r)} b'_n(u_1,X^n_{u_1})\cdots b''_n(u_r,X^n_{u_r})$$
$$\times b'_n(v_1,X^n_{v_1})\cdots b'_n(v_{m_2},X^n_{v_{m_2}})b'_n(u_{r+1},X^n_{u_{r+1}})\cdots b'_n(u_{m_1},X^n_{u_{m_1}})dv_1\cdots dv_{m_2}du_1\cdots du_{m_1}.$$
$$=: I^n_1+I^n_2,$$

where I_1^n and I_2^n denote respectively the two summands in the expression. Denote by $\|\cdot\|$ the maximum norm on $\mathbb{R}^{d\times d\times d}$. Then Minkowski's inequality gives

$$E\|D_{s_2}D_{s_1}X_t^n\|^p = E\|I_1^n + I_2^n\|^p \leq 2^{p-1}\left(E\|I_1^n\|^p + E\|I_2^n\|^p\right)$$

Let $p \geq 1$ and choose $p_1, p_2 \in [1,\infty)$ such that $pp_1 = 2^q$ for some integer q and $\frac{1}{p_1} + \frac{1}{p_2} = 1$. We focus now on the term I_2^n. Then by Girsanov's theorem we have

$$E\|I_2^n\|^p = E\Bigg[\Bigg\| \sum_{m_1\geq 1}\sum_{r=1}^{m_1}\sum_{m_2\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\int_{\Lambda_{m_2}(s_2,u_r)} b_n'(u_1,B_{u_1}^x)\cdots b_n''(u_r,B_{u_r}^x)$$
$$\times b_n'(v_1,B_{v_1}^x)\cdots b_n'(v_{m_2},B_{v_{m_2}}^x)b_n(u_{r+1},B_{u_{r+1}}^x)\cdots b_n'(u_{m_1},B_{u_{m_1}}^x)dv_1\cdots dv_{m_2}du_1\cdots du_{m_1}\Bigg\|^p$$
$$\times\mathcal{E}\left(\sum_{i=1}^d\int_0^T b_n^{(i)}(u,B_u^x)dB_u^{(i)}\right)\Bigg]$$

where $B_t^x := x + B_t$, $t\in[0,T]$.

Then choose $p_2 = 1+\varepsilon$ and $p_1 = \frac{1+\varepsilon}{\varepsilon}$ with $\varepsilon > 0$ sufficiently small and apply Lemma 2.8 to obtain

$$E\|I_2^n\|^p \leq C_\varepsilon\Bigg\| \sum_{m_1\geq 1}\sum_{r=1}^{m_1}\sum_{m_2\geq 1}\int_{\Lambda_{m_1}(s_2,t)}\int_{\Lambda_{m_2}(s_2,u_r)} b_n'(u_1,B_{u_1}^x)\cdots b_n''(u_r,B_{u_r}^x)b_n'(v_1,B_{v_1}^x)$$
$$\times\cdots b_n'(v_{m_2},B_{v_{m_2}}^x)b_n(u_{r+1},B_{u_{r+1}}^x)\cdots b_n'(u_{m_1},B_{u_{m_1}}^x)dv_1\cdots dv_{m_2}du_1\cdots du_{m_1}\Bigg\|^p_{L^{2^q}(\Omega;\mathbb{R}^{d\times d\times d})} \tag{2.19}$$

Now we carry out the product of linear and bilinear forms in the integrand of (2.19). Recall that $b''(u,B_u^x) = \left(\frac{\partial^2}{\partial x_j\partial x_k}b^{(i)}(u,B_u^x)\right)_{i,j,k=1,\dots,d}$ and $b'(u,B_u^x) = \left(\frac{\partial}{\partial x_j}b^{(i)}(u,B_u^x)\right)_{i,j=1,\dots,d}$ where the superscript $b^{(i)}(u,B_u^x)$ here denotes the i-th component of the vector $b(u,B_u^x)$ and $\frac{\partial}{\partial x_j}$, resp. $\frac{\partial^2}{\partial x_j\partial x_k}$, denote the weak derivative of $b^{(i)}(u,B_u^x)$ with respect to the j-th space component, resp. with respect to the j-th and k-th space components. So we represent the second order derivatives as a matrix of matrices in this case, i.e. $b''(t,x) = \nabla\otimes\nabla b(t,x)$ where $\otimes$ denotes the Kronecker tensor product.

Hence we can represent the second order derivatives in the integrand in (2.19) in this manner

$$b''(u, B_u^x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \begin{pmatrix} \frac{\partial}{\partial x_1} b^{(1)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_d} b^{(1)}(u, B_u^x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) \end{pmatrix} \\ \vdots \\ \frac{\partial}{\partial x_d} \begin{pmatrix} \frac{\partial}{\partial x_1} b^{(1)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_d} b^{(1)}(u, B_u^x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) \end{pmatrix} \end{pmatrix} \tag{2.20}$$

The product of $b''(B_u)$ with $b'(B_v)$ is then

$$b''(u, B_u^x) b'(v, B_v^x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \begin{pmatrix} \frac{\partial}{\partial x_1} b^{(1)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_d} b^{(1)}(u, B_u^x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) \end{pmatrix} b'(v, B_v^x) \\ \vdots \\ \frac{\partial}{\partial x_d} \begin{pmatrix} \frac{\partial}{\partial x_1} b^{(1)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_d} b^{(1)}(u, B_u^x) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) & \cdots & \frac{\partial}{\partial x_1} b^{(d)}(u, B_u^x) \end{pmatrix} b'(v, B_v^x) \end{pmatrix} \tag{2.21}$$

As a result

$$b''(u, B_u^x) b'(v, B_v^x) = \left(\sum_{l=1}^{d} \frac{\partial^2}{\partial x_k \partial x_l} b^{(i)}(u, B_u^x) \frac{\partial}{\partial x_j} b^{(l)}(v, B_v^x) \right)_{i,j,k=1}^{d}$$

Hence, taking maximum norm over all products

$$\begin{aligned} E\|I_2^n\|^p \le C_p \Bigg(& \sum_{m_1 \ge 1} \sum_{r=1}^{m_1} \sum_{m_2 \ge 1} \sum_{i,j,k=1}^{d} \sum_{l_1,\dots,l_{m_1+m_2-1}=1}^{d} \Bigg\| \int_{\Lambda_{m_1}(s_2,t)} \int_{\Lambda_{m_2}(s_2,u_r)} \frac{\partial}{\partial x_{l_1}} b_n^{(i)}(u_1, B_{u_1}^x) \\ & \times \frac{\partial}{\partial x_{l_2}} b_n^{(l_1)}(u_2, B_{u_2}^x) \cdots \frac{\partial}{\partial x_{l_{r-1}}} b_n^{(l_{r-2})}(u_{r-1}, B_{u_{r-1}}^x) \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_{l_r}} b_n^{(l_{r-1})}(u_r, B_{u_r}^x) \\ & \times \frac{\partial}{\partial x_{l_{r+1}}} b_n^{(l_r)}(v_1, B_{v_1}^x) \cdots \frac{\partial}{\partial x_{l_{r+m_2}}} b_n^{(l_{r+m_2-1})}(v_{m_2}, B_{v_{m_2}}^x) \frac{\partial}{\partial x_{l_{r+m_2+1}}} b_n^{(l_{r+m_2})}(u_{r+1}, B_{u_{r+1}}^x) \\ & \times \cdots \frac{\partial}{\partial x_j} b_n^{(l_{m_1+m_2-1})}(u_{m_1}, B_{u_{m_1}}^x) dv_1 \cdots dv_{m_2} du_1 \cdots du_{m_1} \Bigg\|_{L^{2^q}(\Omega;\mathbb{R})} \Bigg)^p . \end{aligned} \tag{2.22}$$

Observe the second order partial derivatives in the integrand.

The following step is to apply expectation and get rid of the second order derivatives. To do so, we will use the estimate from Proposition 2.3.

Before applying Proposition 2.3 we need to make the following observation on the integrating regions in connection to (2.22): the iterated integrals of (2.22) can be split up as a sum of integrals where the regions which we integrate over are ordered. Indeed, using formula (2.5) we express the term in (2.22) as follows

$$
\begin{aligned}
&E\|I_2^n\|^p \\
&\leq C_p \left(\sum_{m_1\geq 1} \sum_{r=1}^{m_1} \sum_{m_2 \geq 1} \sum_{i,j,k=1}^{d} \sum_{l_1,\dots,l_{m_1+m_2-1}=1}^{d} \sum_{\sigma\in S_r(m,n)} \left\| \int_{\Lambda_{m_1+m_2}(s_2,t)} \frac{\partial}{\partial x_{l_1}} b_n^{(i)}(w_{\sigma(1)}, B^x_{w_{\sigma(1)}}) \right.\right. \\
&\quad \times \frac{\partial}{\partial x_{l_2}} b_n^{(l_1)}(w_{\sigma(2)}, B^x_{w_{\sigma(2)}}) \cdots \frac{\partial}{\partial x_{l_{r-1}}} b_n^{(l_{r-2})}(w_{\sigma(r-1)}, B^x_{w_{\sigma(r-1)}}) \frac{\partial}{\partial x_k}\frac{\partial}{\partial x_{l_r}} b_n^{(l_{r-1})}(w_{\sigma(r)}, B^x_{w_{\sigma(r)}}) \\
&\quad \times \frac{\partial}{\partial x_{l_{r+1}}} b_n^{(l_r)}(w_{\sigma(r+1)}, B^x_{w_{\sigma(r+1)}}) \cdots \frac{\partial}{\partial x_{l_{r+m_2}}} b_n^{(l_{r+m_2-1})}(w_{\sigma(r+m_2)}, B^x_{w_{\sigma(r+m_2)}}) \\
&\quad \times \frac{\partial}{\partial x_{l_{r+m_2+1}}} b_n^{(l_{r+m_2})}(w_{\sigma(r+m_2+1)}, B^x_{w_{\sigma(r+m_2+1)}}) \cdots \\
&\quad \left.\left. \cdots \frac{\partial}{\partial x_j} b_n^{(l_{m_1+m_2-1})}(w_{\sigma(m_1+m_2)}, B^x_{w_{\sigma(m_1+m_2)}}) dw_{\sigma(1)} \cdots dw_{\sigma(m_1+m_2)} \right\|_{L^{2^q}(\Omega;\mathbb{R})} \right)^p .
\end{aligned} \tag{2.23}
$$

Now that the sets over which we integrate are symmetric we can use deterministic integration by parts or formula (2.4) iteratively in order to write the integrals in (2.23) to the power two as a sum of at most 2^{2m} summands of the form

$$\int_{\Lambda^s_{2m}(s_2,t)} g_1(w_1)\cdots g_{2m}(w_{2m}) dw_1 \cdots dw_{2m}$$

where $m := m_1 + m_2$ and $g_l \in \left\{ \frac{\partial}{\partial x_j} b^{(i)}(\cdot, B^x_\cdot), \frac{\partial^2}{\partial x_{l_1}\partial x_{l_2}} b^{(k)}(\cdot, B^x_\cdot), i,j,k,l_1,l_2 = 1,\dots d \right\}$ and $l = 1,\dots,2m$. Once more, we can write the integrals to the power four as a sum of at most 2^{8m} summands of the form

$$\int_{\Lambda^s_{4m}(s_2,t)} g_1(w_1)\cdots g_{4m}(w_{4m}) dw_1 \cdots dw_{4m}.$$

Repeating this principle, one can write the integrals to the power 2^q as a sum of at most $2^{q2^q m}$ summands of the form

$$\int_{\Lambda^s_{2^q m}(s_2,t)} g_1(w_1)\cdots g_{2^q m}(w_{2^q m}) dw_1 \cdots dw_{2^q m}.$$

Combining this with Proposition 2.3 we obtain

$$
\begin{aligned}
&\Big\| \int_{\Lambda_{m_1+m_2}(s_2,t)} \frac{\partial}{\partial x_{l_1}} b_n^{(i)}(w_{\sigma(1)}, B^x_{w_{\sigma(1)}}) \\
&\times \frac{\partial}{\partial x_{l_2}} b_n^{(l_1)}(w_{\sigma(2)}, B^x_{w_{\sigma(2)}}) \cdots \frac{\partial}{\partial x_{l_{r-1}}} b_n^{(l_{r-2})}(w_{\sigma(r-1)}, B^x_{w_{\sigma(r-1)}}) \frac{\partial}{\partial x_k}\frac{\partial}{\partial x_{l_r}} b_n^{(l_{r-1})}(w_{\sigma(r)}, B^x_{w_{\sigma(r)}}) \\
&\times \frac{\partial}{\partial x_{l_{r+1}}} b_n^{(l_r)}(w_{\sigma(r+1)}, B^x_{w_{\sigma(r+1)}}) \cdots \frac{\partial}{\partial x_{l_{r+m_2}}} b_n^{(l_{r+m_2-1})}(w_{\sigma(r+m_2)}, B^x_{w_{\sigma(r+m_2)}}) \\
&\times \frac{\partial}{\partial x_{l_{r+m_2+1}}} b_n^{(l_{r+m_2})}(w_{\sigma(r+m_2+1)}, B^x_{w_{\sigma(r+m_2+1)}}) \cdots \\
&\cdots \frac{\partial}{\partial x_j} b_n^{(l_{m_1+m_2-1})}(w_{\sigma(m_1+m_2)}, B^x_{w_{\sigma(m_1+m_2)}}) dw_{\sigma(1)} \cdots dw_{\sigma(m_1+m_2)} \Big\|_{L^{2^q}(\Omega;\mathbb{R})} \\
&\le \left(\frac{2^{q2^q(m_1+m_2)} C_{d,p,T}^{2^q(m_1+m_2)} \|b_n'\|_\infty^{2^q(m_1+m_2)} |t-s_2|^{2^{q-1}(m_1+m_2)}}{\Gamma\left(2^{q-1}(m_1+m_2)+1\right)} \right)^{2^{-q}} \\
&= \frac{2^{q(m_1+m_2)} C_{d,p,T}^{m_1+m_2} \|b_n'\|_\infty^{m_1+m_2} |t-s_2|^{(m_1+m_2)/2}}{\left[(2^{q-1}(m_1+m_2))!\right]^{2^{-q}}}
\end{aligned}
\tag{2.24}
$$

Using the bound in (2.24) we get

$$
\begin{aligned}
&E\|I_2^n\|^p \le \\
&\le \left(\sum_{m_1\ge 1} \sum_{r=1}^{m_1} \sum_{m_2\ge 1} C^{m_1+m_2} \frac{d^{m_1+m_2+2} 2^{q(m_1+m_2)} C_{d,p,T}^{m_1+m_2} \|b_n'\|^{m_1+m_2} |t-s_2|^{(m_1+m_2)/2}}{\left[(2^{q-1}(m_1+m_2))!\right]^{2^{-q}}} \right)^p \\
&\le \left(\sum_{m_1\ge 1} \sum_{m_2\ge 1} m_1 C^{m_1+m_2} \frac{d^{m_1+m_2+2} 2^{q(m_1+m_2)} C_{d,p,T}^{m_1+m_2} \|b_n'\|^{m_1+m_2} |t-s_2|^{(m_1+m_2)/2}}{\left[(2^{q-1}(m_1+m_2))!\right]^{2^{-q}}} \right)^p \\
&\le \left(\sum_{m\ge 1} m C^m \frac{d^{m+2} 2^{qm} C_{d,p,T}^m \|b_n'\|_\infty^m |t-s_2|^{m/2}}{\left[(2^{q-1}m)!\right]^{2^{-q}}} \right)^p \\
&\le C_{d,p,T} f(\|b_n'\|_\infty)
\end{aligned}
$$

for some continuous function f only depending on d, p and T. As a result,

$$
\sup_{n\ge 0} \sup_{s_1,s_2\in[0,T]} E\|I_2^n\|^p \le C_{d,p,T} \sup_{n\ge 0} \sup_{s_1,s_2\in[0,T]} f(\|b_n'\|_\infty) < \infty.
$$

Finally, one can bound $E\|I_1^n\|^p$ using exactly the same steps as for I_2^n. □

We are now in a position to state one of the main results of this section on the Malliavin regularity of the solution to SDE (2.6).

Theorem 2.7. *Assume that b satisfies condition (H) for some $k \ge 1$. Let X_t, $t \in [0,T]$ denote the solution to equation (2.6). Then*

$$
X_t \in \bigcap_{p>1} \mathbb{D}^{k+1,p}(\Omega).
$$

Proof. The proof of this more general result relies on Theorem 2.4 by iterating all arguments up to $k+1$. Similarly as before, let $\{b_n\}_{n\geq 1} \subset C^{k+1}(\mathbb{R}^d)$ be an approximating sequence of functions such that $b_n \to b$ a.e. in $t \in [0,T]$ and $x \in \mathbb{R}^d$ as $n \to \infty$ and $\sup_{n\geq 0} |b_n(t,x)| \leq C(1+|x|)$ and $\sup_n \|b_n^{(j)}\| < \infty$, $j \leq k$ where $\|\cdot\|$ denotes the maximum norm in $\mathbb{R}^d \times \overset{j)}{\cdots} \times \mathbb{R}^d$, $j \leq k$. For each $t \in [0,T]$, denote by X_t^n the sequence of random variables in $L^p(\Omega)$ solution to equation (2.6) with drift coefficient b_n. Then we wish to compute the Malliavin derivative of X_t^n upto order $k+1$. This becomes a large expression where the terms increase at a binomial speed. We saw in the proof of Proposition (2.4) that the second order Malliavin derivative of X_t^n can be written as $D_{s_2}D_{s_1}X_t^n = I_1^n + I_2^n$ where the integrals in I_2^n are doubled. If we fix $s_3 \in [0,t]$ then $D_{s_3}D_{s_2}D_{s_1}X_t^n = D_{s_3}I_1^n + D_{s_3}I_2^n = I_1^n + I_2^n + I_3^n + I_4^n$ and so on. Each term I_i^n, $i = 1,2,3,4$ is a sum of integrals of the form (2.18) with at most one factor $b_n^{(3)}$. Iterating this argument, we have that for fixed $0 \leq s_1 \leq \cdots \leq s_{k+1} \leq t$

$$D_{s_{k+1}} \cdots D_{s_1} X_t^n = I_1^n + \cdots + I_{2^k}^n$$

where each I_i^n, $i = 1,2,\ldots,2^k$ is an integral over at most $\Lambda_{m_1+\cdots+m_{k+1}}$ with at most one factor $b_n^{(k+1)}$ and the rest $b_n^{(j)}$, $j \leq k$. This can be readily checked by looking at expression (2.18). Then, estimating $I_{2^k}^n$ implies that all former terms are also bounded.

To illustrate $I_{2^k}^n$ we use expression (2.18) and apply $D_{s_3} \cdots D_{s_{k+1}}$ and focus on the last term. In order to simplify notation and make the reading clearer we consider indices $m_1,\ldots,m_{k+1}$, $r_1,\ldots,r_k \in \mathbb{N} \setminus \{0\}$ and denote

$$\sum_{\substack{m_1,\ldots,m_{k+1}\\ r_1,\ldots,r_k}} := \sum_{m_1\geq 1} \sum_{r_1=1}^{m_1} \sum_{m_2\geq 1} \sum_{r_2=1}^{m_1+m_2} \cdots \sum_{r_k=1}^{m_1+\cdots+m_k} \sum_{m_{k+1}\geq 0},$$

as well as,

$$\int_\Delta := \int_{\Lambda_{m_1}(s_{k+1},t)} \int_{\Lambda_{m_2}(s_{k+1},t)} \cdots \int_{\Lambda_{m_{k+1}}(s_{k+1},t)}.$$

Then $I_{2^k}^n$ will take the following form

$$I_{2^k}^n = \sum_{\substack{m_1,\ldots,m_{k+1}\\ r_1,\ldots,r_k}} \int_\Delta \mathcal{A}(u_1^1,\ldots,u_{m_1}^1,\ldots,u_1^{k+1},\ldots,u_{m_{k+1}}^{k+1}) du_1^{k+1} \cdots du_{m_{k+1}}^{k+1} \cdots du_1^1 \cdots du_{m_1}^1$$

with integrand

$$\mathcal{A} := g_n(u_1^1) \cdots g_n(u_{r_1}^1) \Bigg[g_n(u_1^2) \cdot g_n(u_{r_2}^2) \Bigg[\cdots g_n(u_1^{k+1}) \cdots$$
$$\cdots g_n(u_{m_{k+1}}^{k+1}) \Bigg] g_n(u_{r_k+1}^k) \cdots g_n(u_{m_2}^2) \Bigg] g_n(u_{r_1+1}^1) \cdots g_n(u_{m_1}^1)$$

where the functions g_n denote an element in the set

$$g_n \in \{Db_n, D^2b_n, \cdots, D^{k+1}b_n\}.$$

Then, using exactly the same procedure as for I_1^n and I_2^n, mutatis mutandis, we obtain an integral of products of partial derivatives of at most order $k+1$, this together with Proposition 2.3 one is able to get rid of the $k+1$-th derivative as we did for I_2^n in Theorem 2.7. □

To emphasize that the solution depends on the initial point x we write X_t^x. Next result gives a criteria for the regularity of $x \mapsto X_t^x$ in the space variable in the Sobolev sense.

Theorem 2.8. *Assume that b satisfies condition (H) for some $k \geq 1$. Let $U \subset \mathbb{R}^d$ be an open bounded set and X_t, $t \in [0,T]$ denote the solution to equation (2.6). Then*

$$X_t^{\cdot} \in \bigcap_{p>1} L^2\big(\Omega, W^{k+1,p}(U)\big).$$

Proof. This result actually follows by observing that the process $\frac{\partial}{\partial x} X_t^x$ satisfies the following linear ODE

$$\frac{\partial}{\partial x} X_t^x = \mathcal{I}_d + \int_0^t b'(u, X_u^x) \frac{\partial}{\partial x} X_u^x du.$$

This equation is the same as (2.15) when $s=0$. Using this observation, in connection with the same method employed in the proof of Theorem 2.7 by replacing the Malliavin derivative of X_t with $\frac{\partial}{\partial x} X_t^x$ we get that for the approximating sequence of solutions $X_t^{n,x}$, $n \geq 0$ described in Theorem 2.7 we have

$$\sup_{n\geq 0} \sup_{x \in U} E\left[\| \frac{\partial^j}{\partial x^j} X_t^{n,x} \|^p \right] < \infty$$

for all $j = 0, \ldots, k+1$ and any $p > 1$ so $X_t^{n,\cdot}$ is bounded in the Sobolev norm $L^2(\Omega, W^{k+1,p}(U))$ for each $n \geq 0$. Indeed

$$\begin{aligned}
\sup_{n\geq 0} \|X_t^{n,\cdot}\|^2_{L^2(\Omega, W^{k+1,p}(U))} &= \sup_{n\geq 0} \sum_{i=0}^{k+1} E\left[\| \frac{\partial^i}{\partial x^i} X_t^{n,\cdot} \|^2_{L^p(U)} \right] \\
&\leq \sum_{i=0}^{k+1} \int_U \sup_{n\geq 0} E\left[\| \frac{\partial^i}{\partial x^i} X_t^{n,x} \|^p \right] dx \\
&< \infty.
\end{aligned}$$

Since $L^2(\Omega, W^{k+1,p}(U))$ is reflexive for $p > 1$, by Banach-Alaoglu's theorem we get that the set $\{X_t^{n,x}\}_{n\geq 0}$ is weakly compact in the $L^2(\Omega, W^{k+1,p}(U))$ topology. Thus, there exists a subsequence $n(j)$, $j \geq 0$ such that

$$X_t^{n(j),\cdot} \xrightarrow[j\to\infty]{w} Y \in L^2(\Omega, W^{k+1,p}(U)).$$

On the other hand, we have that $X_t^{n,x} \to X_t^x$ strongly in $L^p(\Omega)$, so by uniqueness of the limit we can conclude that

$$X_t^{\cdot} = Y, \quad P - a.s.$$

□

Remark 2.9. *The previous result actually gives classical derivatives of the solution upto order $k+\alpha$ with $\alpha \in (0,1)$ as a consequence of the Sobolev embedding (2.3).*

2.4 Application to the regularity of densities

As mentioned in the introduction one implication of improving the Malliavin regularity of SDEs with drift coefficient satisfying hypotheses **(H)** is that the finite dimensional laws have k-times differentiable densities due to a result by V.Bally and L.Caramellino, see [5]. We see this as an improvement of the regularity criterion given by [73] for the additive noise case. In addition, we see that the boundedness of b is not needed.

The following is a consequence of Theorem 2.7 for the special case $d = 1$ and illustrates how we may gain regularity of the densities of solutions to (2.6) and provide with an explicit expression for the density and its derivatives. Later on, we will show it for higher dimensions.

Corollary 2.10. *For $d = 1$, let p_{X_t} denote the density of the solution X_t to equation (2.6) for a given $t \in [0, T]$. If b satisfies (H) for some $k \geq 1$ then $p_{X_t} \in C^{k-1}(\mathbb{R})$ for every $t \in [0, T]$.*

Proof. Let $G_0, G_1, G_2, \dots, G_k$ be the random variables defined as $G_0 = 1$ and for each $i = 1, \dots, k$

$$G_i = \delta\left(G_{i-1} \cdot \left(\int_0^T D_s X_t ds\right)^{-1}\right).$$

It is known that if $X_t \in \mathbb{D}^{1,2}(\Omega)$, $\int_0^T D_s X_t ds \neq 0$, $P-a.s.$ and $G_i \left(\int_0^T D_s X_t ds\right)^{-1} \in \mathrm{Dom}(\delta)$ for each $i = 0, \dots, k$ then X_t has a density of class $C^k(\mathbb{R})$ and

$$\frac{d^i}{dy^i} p_{X_t}(y) = (-1)^i E\left[\mathbf{1}_{\{X_t > y\}} G_{i+1}\right] \tag{2.25}$$

for each $i = 0, \dots, k$. See [90, p115].

We will then prove that $G_i \left(\int_0^T D_s X_t ds\right)^{-1} \in \mathrm{Dom}(\delta)$ for $i = 0, \dots, k-1$. First, observe that for dimension $d = 1$ we can easily solve the linear SDE for $D_s X_t$ and write

$$D_s X_t = \exp\left\{\int_s^t b'(u, X_u) du\right\} \tag{2.26}$$

where b' denotes the weak derivative of b (one may also use local time to express (2.26) independently of b' if b is non-regular, see [39]). Hence, for any $t \in [0, T]$ there is an $\varepsilon > 0$ such that $\int_0^T D_s X_t ds \geq \varepsilon > 0$. Since $x \mapsto \frac{1}{x}$ is smooth on the domain (ε, ∞) we see that $\left(\int_0^T D_s X_t ds\right)^{-1} \in \mathbb{D}^{k,2}(\Omega)$ since $X_t \in \mathbb{D}^{k+1,2}(\Omega)$ by Theorem 2.7.

Denote $F := \left(\int_0^T D_s X_t ds\right)^{-1}$. Now, since $F \in \mathbb{D}^{1,2}(\Omega)$ we have $F \in \mathrm{Dom}(\delta)$ and $G_1 = \delta(F) = FW(T) + \int_0^T D_s F ds$. Then we see that $G_1 \in \mathbb{D}^{1,2}(\Omega)$ and hence $G_1 F \in \mathbb{D}^{1,2}(\Omega)$ therefore $G_1 F \in \mathrm{Dom}(\delta)$ with $G_2 = \delta(G_1 F) = G_1 F W(T) - \int_0^T [D_s G_1 F + G_1 D_s F] ds$. Again, it is readily checked that $G_2 \in \mathbb{D}^{1,2}(\Omega)$ since $G_1, F \in \mathbb{D}^{1,2}(\Omega)$ so $G_2 F \in \mathrm{Dom}(\delta)$. For a fixed $i = 0, \dots, k-1$ we have $G_i, F \in \mathbb{D}^{1,2}(\Omega)$ therefore $G_i F \in \mathrm{Dom}(\delta)$ with $G_{i+1} = \delta(G_i F) = G_i F W(T) - \int_0^T [D_s G_i F + G_i D_s F] ds$. So G_i is well-defined for $i = 0, \dots, k$ but we can not say anything about G_{k+1} so p_{X_t} is at least $k-1$-times differentiable with derivatives given by (2.25). $\square$

As a consequence of the Malliavin regularity we have shown for SDEs of the form (2.6) we may apply the results by V.Bally and L.Caramellino, see [5], to be able to obtain regularity of the densities, also in higher dimension. In order to do so, we need to study integrability properties of the Malliavin covariance matrix. Let us denote

$$\gamma_{X_t}^{ij} := \langle D.X_t^{(i)}, D.X_t^{(j)} \rangle_{L^2([0,T])}, \quad i,j = 1,\ldots,d,$$

the Malliavin covariance matrix of the process X_t, given $t \in [0,T]$. We will say that $\gamma_{X_t} = (\gamma_{X_t}^{ij})_{i,j=1,\ldots,d}$ satisfies the *non-degeneracy condition* whenever

$$(\det \gamma_{X_t})^{-1} \in \bigcap_{p\geq 1} L^p(\Omega). \tag{2.27}$$

Next, we invoke a result by [5, Proposition 23] which gives us the desired properties on the density of X_t, $t \in [0,T]$.

Proposition 2.11. *Let $F = (F^1,\ldots,F^d)$ with $F^1,\ldots,F^d \in \bigcap_{p\geq 1} \mathbb{D}^{k+1,p}(\Omega)$. Assume that condition (2.27) holds for γ_F. Denote by p_F the density of F. Then $p_F \in C^{k-1,\alpha}(\mathbb{R}^d)$ with $\alpha < 1$, i.e. p_F is $k-1$-times differentiable with Hölder continuous derivatives of exponent $\alpha < 1$.*

In view of the above result we only need to check that the non-degeneracy condition (2.27) is fulfilled. To do so, we use the following intermediate result.

Lemma 2.12. *Let $Z : \Omega \to E$ be a random variable taking values on a separable Banach space with norm $\|\cdot\|_E$. Fix $p > 0$. Then the following are equivalent*

(i)

$$E\left[\|Z\|_E^{-p}\right] < \infty. \tag{2.28}$$

(ii) There exists $\varepsilon_0 > 0$, depending on p, such that

$$\int_0^{\varepsilon_0} \varepsilon^{-(p+1)} P(\|Z\|_E^2 < \varepsilon) d\varepsilon < \infty.$$

Proof. We have that, for any positive integrable random variable Y,

$$E[Y] = \int_0^\infty P(Y > \eta) d\eta.$$

Condition (i) implies that $\|Z\|_E > 0$ P-a.s. so

$$\begin{aligned} E\left[\|Z\|_E^{-2p}\right] &= \int_0^{\eta_0} P(\|Z\|_E^{-2p} > \eta) d\eta + \int_{\eta_0}^\infty P(\|Z\|_E^{-2p} > \eta) d\eta \\ &\leq \eta_0 + \int_{\eta_0}^\infty P(\|Z\|_E^{-2p} > \eta) d\eta \\ &= \eta_0 + p \int_0^{\eta_0^{-1/p}} \varepsilon^{-(p+1)} P(\|Z\|_E^2 < \varepsilon) d\varepsilon \end{aligned}$$

where in the last step we have used the change of variables $\eta = \varepsilon^{-p}$. □

Now we are in a position to prove the non-degeneracy condition for the Malliavin matrix associated to the solution of the SDE (2.6). The proof of this result is much inspired in Proposition 8.1 from [103].

Proposition 2.13. *Let X_t, $t \in [0,T]$ be the solution to SDE (2.6) with drift coefficient b satisfying condition (H) for $k = 1$. Then the Malliavin covariance matrix γ_{X_t} satisfies*

$$(\det \gamma_{X_t})^{-1} \in \bigcap_{p\geq 1} L^p(\Omega)$$

Proof. Consider X_t^n with drift coefficient b_n approximating b a.e. such that $\sup_{n\geq 0} \|b_n'\|_\infty < \infty$.

It suffices to show that

$$\sup_{n\geq 0} E\left[\left| \int_0^T \|D_s X_t^n\|_\infty^2 ds \right|^{-p} \right] < \infty$$

for any $p \geq 1$.

Recall that for $0 \leq s \leq t$, $t \in [0,T]$ we have

$$D_s X_t^n = \mathcal{I}_d + \sum_{m\geq 1} \int_{s<u_1<\cdots<u_m<t} b_n'(u_1, X_{u_1}^n) \cdots b_n'(u_m, X_{u_m}^n) du_1 \cdots du_m.$$

Then for any $\delta > 0$, $t - \delta > 0$ one has

$$\int_0^T \|D_s X_t^n\|_\infty^2 ds \geq \int_{t-\delta}^t \|D_s X_t^n\|_\infty^2 ds \geq \frac{\delta}{2} - I_n(t,\delta)$$

where

$$I_n(t,\delta) := \int_{t-\delta}^t \left\| \sum_{m\geq 1} \int_{s<u_1<\cdots<u_m<t} b_n'(u, X_{u_1}^n) \cdots b_n'(u, X_{u_m}^n) du_1 \cdots du_m \right\|_\infty^2 ds.$$

Clearly, we have

$$\sup_{n\geq 0} E\left[|I_n(t,\delta)|^p \right] \leq C\delta^p \tag{2.29}$$

since b_n', $n \geq 0$ are uniformly bounded.

Then by the previous estimates

$$\begin{aligned} P\left(\|D.X_t^n\|_{L^2(\Omega,\mathbb{R}^{d\times d})}^2 < \varepsilon \right) &\leq P\left(\int_{t-\delta}^t \|D_s X_t^n\|_\infty^2 ds < \varepsilon \right) \\ &\leq P\left(I_n(t,\delta) \geq \frac{\delta}{2} - \varepsilon \right) \\ &\leq \left(\frac{\delta}{2} - \varepsilon \right)^{-p} E[|I_n(t,\delta)|^p] \end{aligned}$$

for any $p \geq 1$ due to Chebyshev's inequality. Now, by estimate (2.29) we obtain that

$$\sup_{n\geq 0} P\left(\|D.X_t^n\|^2_{L^2(\Omega,\mathbb{R}^{d\times d})} < \varepsilon\right) \leq C\left(\frac{\delta}{2}-\varepsilon\right)^{-p}\delta^p.$$

By virtue of Lemma 2.12 we can conclude if we find $\delta : (0,\infty) \to \mathbb{R}$, $\varepsilon \mapsto \delta(\varepsilon)$ such that $\lim_{\varepsilon\searrow 0}\delta(\varepsilon) = 0$ and

$$\int_0 \varepsilon^{-(p+1)}\left(\frac{\delta(\varepsilon)}{2}-\varepsilon\right)^{-p}\delta(\varepsilon)^p d\varepsilon < \infty$$

for an arbitrary large $p \geq 1$.

We claim that

$$\delta(\varepsilon) := \left|\frac{2\varepsilon^{\frac{1}{2p}+2}}{\varepsilon^{\frac{1}{2p}+1}-2}\right|$$

does the job. □

Finally, we are able to state our criteria to determine the regularity of densities of solutions to SDEs.

Corollary 2.14. *Let X_t^x, $t \in [0,T]$ be the strong solution to SDE (2.6). Assume b satisfies condition (H) for some integer $k \geq 1$. Then the density p_{X_t} belongs to $C^{k-1,\alpha}(\mathbb{R}^d)$, $\alpha < 1$, i.e. $k-1$-times continuously differentiable with Hölder continuous derivatives with exponent $\alpha < 1$.*

We end this section by giving an example that shows that the Malliavin regularity we obtained in Theorem 2.7 is optimal when $k = 1$, for the general criteria we conjecture it is also optimal.

Example 2.15. *In this example we show that Theorem 2.4 is an optimal result in the sense that, if b is of linear growth and one time weakly differentiable with bounded derivative then $X_t \in \mathbb{D}^{2,p}(\Omega)$ for all $p \geq 1$ and $X_t \notin \mathbb{D}^{3,p}(\Omega)$ for any $p \geq 1$. Just choose b, in dimension $d = 1$, to be such that*

$$b'(x) = \mathbf{1}_{(0,\infty)}(x),\ x \in \mathbb{R}.$$

Then fix $t \in [0,T]$ and for $s_1 \leq t$

$$D_{s_1}X_t = \exp\left\{\int_{s_1}^t b'(X_u)du\right\}.$$

Denote by $\tilde{b}(x) := b(a) + \int_a^x b(y)dy$, $a \in \mathbb{R}$ a primitive of b. Itô's formula implies

$$D_{s_1}X_t = \exp\left\{2\tilde{b}(X_t) - 2\tilde{b}(X_{s_1}) - 2\int_{s_1}^t b^2(X_u)du - 2\int_{s_1}^t b(X_u)dB_u\right\}.$$

Then by Theorem 2.7, $D_{s_1}X_t \in \mathbb{D}^{1,2}(\Omega)$ for all $t \in [0,T]$. So for $s_2 \leq t$

$$\begin{aligned} D_{s_2}D_{s_1}X_t = D_{s_1}X_t\Bigg(& 2b(X_t)D_{s_2}X_t - 2b(X_{s_1})D_{s_2}X_{s_1}) - 4\int_{s_1\vee s_2}^t b(X_u)b'(X_u)D_{s_2}X_u du \\ & - 2b(X_{s_2}) - 2\int_{s_1\vee s_2}^t b'(X_u)D_{s_2}X_u dB_u)\Bigg). \end{aligned} \tag{2.30}$$

Now observe that $b(X_t) \in \cap_{p\geq 1}\mathbb{D}^{1,p}(\Omega)$ for all $t \in [0,T]$, hence all terms are immediately Malliavin differentiable with all moments except from maybe $\int_{s_1\vee s_2}^t b(X_u)b'(X_u)D_{s_2}X_u du$ and $\int_{s_1\vee s_2}^t b'(X_u)D_{s_2}X_u dB_u$. The stochastic integral is in fact not Malliavin differentiable. Indeed, by [90, Lemma 1.3.4]

$$\int_{s_1\vee s_2}^t b'(X_u)D_{s_2}X_u dB_u \in \mathbb{D}^{1,2}(\Omega) \tag{2.31}$$

if, and only if

$$b'(X_u)D_{s_2}X_u \in \mathbb{D}^{1,2}(\Omega).$$

On the other hand we have $b'(X_u) = \mathbf{1}_{(0,\infty)}(X_u) \notin \mathbb{D}^{1,2}(\Omega)$ since $0 < P(0 < X_u < \infty) < 1$, see [90, Proposition 1.2.6], and so $D_{s_1}X_t\int_{s_1\vee s_2}^t b'(X_u)D_{s_2}X_u dB_u \notin \mathbb{D}^{1,2}(\Omega)$.

Let us finally prove that

$$Y_t := \int_{s_1\vee s_2}^t b(X_u)b'(X_u)D_{s_2}X_u du \in \mathbb{D}^{1,2}(\Omega).$$

Let $\{b_n\}_{\{n\geq 0\}}$ be a sequence of smooth functions such that $b_n(x) \to b(x)$ a.e. in $x \in \mathbb{R}$ as $n \to \infty$ and $\sup_{n\geq 0}\|b_n'\|_\infty < \infty$ and $b_n(x), b_n'(x), b_n''(x) \geq 0$ for all $x \in \mathbb{R}$, we claim that this is trivially possible by the very concrete shape of the function b in this example. Define

$$Y_t^n := \int_{s_1\vee s_2}^t b(X_u)b_n'(X_u)D_{s_2}X_u du.$$

Clearly, $Y_t^n \to Y_t$ in $L^2(\Omega)$ for all $t \in [0,T]$. We only need to bound $\|D.Y_t^n\|_{L^2([0,T]\times\Omega)}$ uniformly in $n \geq 0$. Then

$$\begin{aligned} D_{s_3}Y_t^n = & \int_{s^*}^t b'(X_u)D_{s_3}X_u b_n'(X_u)D_{s_2}X_u du \\ & + \int_{s^*}^t b(X_u)b_n''(X_u)D_{s_3}X_u D_{s_2}X_u du + \int_{s^*}^t b(X_u)b_n'(X_u)D_{s_3}D_{s_2}X_u du \end{aligned}$$

where $s^* := \max\{s_1, s_2, s_3\}$.

Then the critical term is

$$I_n := E\left[\int_0^T \left(\int_{s^*}^t b(X_u)b_n''(X_u)D_{s_3}X_u D_{s_2}X_u du\right)^2 ds_3\right]$$

Denote $\tilde{B}_{s,t} := \exp\left\{\int_s^t b'(B_u^x)du\right\}$*. Then, by Girsanov's theorem and Lemma 2.8 we have for a suitable* $\varepsilon > 0$

$$\begin{aligned} I_n &= \int_0^T E\left[\left(\int_{s^*}^t b(B_u^x)b_n''(B_u^x)\tilde{B}_{s_3,u}\tilde{B}_{s_2,u}du\right)^2 \mathcal{E}\left(\int_0^T b(B_u^x)dB_u\right)\right] ds_3 \\ &\leq C_\varepsilon \int_0^T E\left[\left(\int_{s^*}^t b(B_u^x)b_n''(B_u^x)\tilde{B}_{s_3,u}\tilde{B}_{s_2,u}du\right)^{2\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{1+\varepsilon}} ds_3. \end{aligned}$$

Now we focus on the expectation. Choose $\varepsilon > 0$ *so that* $p := 2\frac{1+\varepsilon}{\varepsilon}$ *is a natural number. Then since* $|\tilde{B}_{s_3,u_i}\tilde{B}_{s_2,u_i}| \leq e^{(|t-s_3|+|t-s_2|)\|b'\|_\infty} \leq C < \infty$ *and since* b *and* b_n'' *are positive we have*

$$\begin{aligned} &\left|E\left[\int_{s^*}^t \cdots \int_{s^*}^t \prod_{i=1}^p b(B_{u_i}^x)b_n''(B_{u_i}^x)\tilde{B}_{s_3,u_i}\tilde{B}_{s_2,u_i}du_1\cdots du_p\right]\right| \leq \\ &\qquad \leq E\left[\int_{s^*}^t \cdots \int_{s^*}^t \prod_{i=1}^p \left|b(B_{u_i}^x)b_n''(B_{u_i}^x)\tilde{B}_{s_3,u_i}\tilde{B}_{s_2,u_i}\right|du_1\cdots du_p\right] \\ &\qquad \leq CE\left[\int_{s^*}^t \cdots \int_{s^*}^t \prod_{i=1}^p \left|b(B_{u_i}^x)b_n''(B_{u_i}^x)\right|du_1\cdots du_p\right] \\ &\qquad = CE\left[\int_{s^*}^t \cdots \int_{s^*}^t \prod_{i=1}^p b(B_{u_i}^x)b_n''(B_{u_i}^x)du_1\cdots du_p\right]. \end{aligned}$$

Then since $(u_1,\ldots,u_p) \mapsto b(B_{u_1}^x)b_n''(B_{u_1}^x)\cdots b(B_{u_p}^x)b_n''(B_{u_p}^x)$ *is symmetric we may write*

$$\begin{aligned} &E\left[\int_{s^*}^t \cdots \int_{s^*}^t \prod_{i=1}^p b(B_{u_i}^x)b_n''(B_{u_i}^x)du_1\cdots du_p\right] \\ &\qquad\qquad \leq p!E\left[\int_{s^*<u_1<\cdots<u_p<t} \prod_{i=1}^p b(B_{u_i}^x)b_n''(B_{u_i}^x)du_1\cdots du_p\right] \end{aligned}$$

and the last may be bounded independently of b_n'' *by using Proposition 2.3. In fact,*

$$\sup_{s^*\in[0,T]}\sup_{n\geq 0} E\left[\int_{s^*<u_1<\cdots<u_p<t} \prod_{i=1}^p b(B_{u_i}^x)b_n''(B_{u_i}^x)du_1\cdots du_p\right] \leq C$$

for a finite constant C*. So*

$$\sup_{n\geq 0} I_n < \infty$$

being thus $Y_t \in \mathbb{D}^{1,2}(\Omega)$ *for every* $t \in [0,T]$*.*

In a summary, we have in (2.30) a sum of Malliavin differentiable terms except for the last one $-4D_{s_1}X_t\int_{s_1\vee s_2}^t b'(X_u)D_{s_2}X_u dB_u$ *. In conclusion* $D_{s_2}D_{s_1}X_t \notin \mathbb{D}^{1,2}(\Omega)$*.*

Finally, we give an extension of Theorem 2.7 and Corollary 2.14 to a class of non-degenerate $d-$dimensional Itô-diffusions.

Theorem 2.16. *Consider the time-homogeneous $\mathbb{R}^d-$valued SDE*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x \in \mathbb{R}^d, \quad 0 \leq t \leq T, \tag{2.32}$$

where the coefficients $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^d \times \mathbb{R}^d$ are Borel measurable. Require that there exists a bijection $\Lambda : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, which is twice continuously differentiable. Let $\Lambda_x : \mathbb{R}^d \longrightarrow L\left(\mathbb{R}^d, \mathbb{R}^d\right)$ and $\Lambda_{xx} : \mathbb{R}^d \longrightarrow L\left(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d\right)$ be the corresponding derivatives of Λ and assume that

$$\Lambda_x(y)\sigma(y) = id_{\mathbb{R}^d} \text{ for } y \text{ a.e.}$$

as well as

$$\Lambda^{-1} \text{ is Lipschitz continuous.}$$

Suppose that the function $b_* : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ given by

$$\begin{aligned} b_*(x) &:= \Lambda_x\left(\Lambda^{-1}(x)\right)\left[b(\Lambda^{-1}(x))\right] \\ &+ \frac{1}{2}\Lambda_{xx}\left(\Lambda^{-1}(x)\right)\left[\sum_{i=1}^d \sigma(\Lambda^{-1}(x))\left[e_i\right], \sum_{i=1}^d \sigma(\Lambda^{-1}(x))\left[e_i\right]\right] \end{aligned}$$

satisfies condition (H), where e_i, $i = 1, \ldots, d$, is a basis of $\mathbb{R}^d$. Then the conclusions of Theorem 2.7 and Corollary 2.14 also apply to X_t, $t \in [0, T]$ and its density.

Proof. The proof can be directly obtained from Itô's Lemma. See [83]. □

2.5 A classical solution to the stochastic transport equation

The Sobolev regularity of the solution shown in Theorem 2.4 with respect to the initial condition entitles us to construct a classical solution to the *stochastic transport equation* when the drift is Lipschitz which to our knowledge is not proved.

The Stochastic Transport Equation is written in differential form

$$\begin{cases} d_t u(t,x) + \nabla u(t,x) \cdot b(t,x)dt + \sum_{i=1}^d e_i \cdot \nabla u(t,x) \circ dB_t^{(i)} = 0 \\ u(0,x) = u_0(x), \end{cases} \tag{2.33}$$

where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a given vector field and $u_0 : \mathbb{R}^d \to \mathbb{R}$ a given initial data. The stochastic integration is understood in the Stratonovich sense.

Definition 2.17 (Classical solution)**.** *Let u_0 and b be given functions. We say that a stochastic process $u \in L^\infty(\Omega \times [0,T] \times \mathbb{R}^d)$ is a classical solution to (2.33) if*

1. *There exists a measurable set $\tilde{\Omega} \subset \Omega$ with full measure such that for fixed $t \in [0,T]$ and $p \geq 1$, the mapping $x \mapsto u(\omega,t,x)$ is in $W^{2,p}_{loc}(\mathbb{R}^d)$ on $\tilde{\Omega}$;*

2. *For fixed $x \in \mathbb{R}^d$ there are $(\mathcal{F}_t)$-adapted versions of $t \mapsto u(t,x)$ and $t \mapsto \nabla u(t,x)$;*

3. *The following integral equation is satisfied*

$$u(t,x) + \int_0^t b(s,x)\cdot\nabla u(s,x)ds + \sum_{i=1}^{d}\int_0^t e_i\cdot\nabla u(s,x)\circ dB_s^{(i)} = u_0(x) \qquad (2.34)$$

for a.e. $(\omega,x)\in\Omega\times\mathbb{R}^d$.

Notice that we are using the Stratonovich integral in our definition, but following the same idea as in [47], Lemma 13, we can recast (2.34) in Itô-form as

$$u(t,x)+\int_0^t b(s,x)\cdot\nabla u(s,x)ds+\sum_{i=1}^{d}\int_0^t e_i\cdot\nabla u(s,x)dB_s^{(i)}+\frac{1}{2}\int_0^t \Delta u(s,x)ds = u_0(x). \quad (2.35)$$

We will use these formulations interchangeably.

Before we proceed further we will introduce the concept of stochastic flow associated to SDE (2.6):

Definition 2.18 (Stochastic flow of diffeomorphisms)**.** *A function* $\varphi:[0,T]\times[0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$, $\varphi_{s,t}(x,\omega)$ *is said to be a stochastic flow of diffeomorphisms of the SDE (2.6) if there exists a full-measure set* $\tilde{\Omega}\in\mathcal{F}$ *such that for any* $\omega\in\tilde{\Omega}$ *the following holds true:*

(i) $\varphi_{s,t}(x,\omega)$, $s,t\in[0,T]$, $x\in\mathbb{R}$ *is a (global) strong solution to the SDE (2.6).*

(ii) $\varphi_{s,t}(x,\omega)$ *is continuous in* $(s,t,x)\in[0,T]\times[0,T]\times\mathbb{R}$.

(iii) $\varphi_{s,t}(\cdot,\omega)=\varphi_{u,t}(\cdot,\omega)\circ\varphi_{s,u}(\cdot,\omega)$ *for any* $s,u,t\in[0,T]$.

(iv) $\varphi_{s,s}(x,\omega)=x$ *for all* $x\in\mathbb{R}$ *and* $s\in[0,T]$.

(v) $\varphi_{s,t}(\cdot,\omega):\mathbb{R}^d\to\mathbb{R}^d$ *are diffeomorphisms (of class* C^k*) for all* $s,t\in[0,T]$.

For the rest of this section we will assume that b satisfies condition (H) for $k=1$ which in particular means that b is globally Lipschitz, uniformly in time.

To get a globally defined (i.e. on the entire $\mathbb{R}^d$) stochastic flow of diffeomorphisms of the SDE (2.6) we notice that since b is uniformly Lipschitz there exists a unique solution to

$$\varphi_{s,t}(x,\omega) = x + \int_s^t b(r,\varphi_{s,r}(x,\omega))dr + B_t(\omega) - B_s(\omega)$$

for *all* $\omega\in\Omega$.

It is easy to check conditions (i) to (iv) in Definition 2.18 holds for all $\omega\in\Omega$.

Fix $p\geq 1$ and $N\in\mathbb{N}$ and invoke Theorem 2.8 to guarantee that there exists a measurable subset $\Omega_N\subset\Omega$ with full measure such that the local solution

$$\varphi_{s,t}^N(x,\omega) = x + \int_s^t b(r,\varphi_{s,r}^N(x,\omega))dr + B_t(\omega) - B_s(\omega).$$

satisfies $\varphi_{s,t}^N(\cdot,\omega)\in W^{2,p}(B(0,N))$ for all $\omega\in\Omega_N$ and $x\in B(0,N)$. By uniqueness we have that $\varphi_{s,t}|_{B(0,N)\times\Omega_N}=\varphi_{s,t}^N$.

If we let $\tilde{\Omega} := \cap_{N=1}^{\infty}\Omega_N$, we get that $\mathbb{R}^d \ni x \mapsto \varphi_{s,t}(x,\omega) \in \mathbb{R}^d$ is twice weakly differentiable for every $\omega \in \tilde{\Omega}$, and thus condition (v) in 2.18 is satisfied for $k = 1$.

In [47] the authors study (2.33) under the considerably weaker condition (at least for $d > 1$), $b \in L^1_{loc}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$, div$b \in L^1_{loc}([0,T] \times \mathbb{R}^d)$ and $u_0 \in L^\infty(\mathbb{R}^d)$. However, in this case, one is restricted to study analytically weak solutions in the sense that for every test function $\theta \in C_0^\infty(\mathbb{R}^d)$ one has

$$\begin{aligned}\int_{\mathbb{R}^d} u(t,x)\theta(x)dx = &\int_{\mathbb{R}^d} u_0(x)\theta(x)dx \\ &+ \int_0^t \int_{\mathbb{R}^d} u(s,x)\left[b(t,x)\cdot\nabla\theta(x) + \mathrm{div}b(t,x)\theta(x)\right]dxds \\ &+ \sum_{i=1}^d \int_0^t \left(\int_{\mathbb{R}^d} u(s,x)\cdot\frac{\partial}{\partial x_i}\theta(x)dx\right)\circ dB_s^{(i)}.\end{aligned} \tag{2.36}$$

Moreover, the equation is uniquely solved by $u(t,x) = u_0(\varphi_t^{-1}(x))$.

Although we consider more restrictive coefficients, we arrive at an analytically stronger solution:

Theorem 2.19. *Let b satisfy condition (H) for $k = 1$ and $u_0 \in C_b^2(\mathbb{R}^d)$. Then there exists a unique classical solution to the stochastic transport equation.*

Moreover, the equation is explicitly solved by $u(t,x) = u_0(\varphi_t^{-1}(x))$.

Proof. By the above discussion we know that for every test function $\theta \in C_0^\infty(\mathbb{R}^d)$, the equation (2.36) is satisfied P-a.s. by $u(t,x) = u_0(\varphi_t^{-1}(x))$. We now choose $\tilde{\Omega}$ such that $x \mapsto \varphi_t^{-1}(x)$ is in $W^{2,p}_{loc}(\mathbb{R}^d)$ on $\tilde{\Omega}$. Then we get that u satisfies condition (i) and (ii) of Definition 2.17, and by integration by parts we have

$$\begin{aligned}\int_{\mathbb{R}^d} u(t,x)\theta(x)dx = &\int_{\mathbb{R}^d} u_0(x)\theta(x)dx \\ &- \int_0^t \int_{\mathbb{R}^d} \nabla u(s,x)\cdot b(t,x)\theta(x)dxds \\ &- \sum_{i=1}^d \int_0^t \left(\int_{\mathbb{R}^d} \nabla u(s,x)\theta(x)dx\right)\circ dB_s^{(i)}.\end{aligned}$$

or equivalently

$$\begin{aligned}\int_{\mathbb{R}^d} u(t,x)\theta(x)dx = &\int_{\mathbb{R}^d} u_0(x)\theta(x)dx \\ &- \int_{\mathbb{R}^d}\int_0^t \nabla u(s,x)\cdot b(t,x)\theta(x)dsdx \\ &- \int_{\mathbb{R}^d}\sum_{i=1}^d \int_0^t \nabla u(s,x)tdB_s^{(i)}\theta(x)dx - \frac{1}{2}\int_{\mathbb{R}^d}\int_0^t \Delta u(s,x)ds\theta(x)dx.\end{aligned}$$

Since θ was arbitrary, this proves the claim. $\square$

Chapter 3

The Bismut-Elworthy-Li formula for mean-field stochastic differential equations

David R. Baños

Abstract: We generalise the so-called Bismut-Elworthy-Li formula to a class of mean-field differential equations whose coefficients might depend on the law of the solution.

3.1 Introduction

It is known that the spatial derivative of the solution to the (backward) Kolmogorov equation can be represented as an expectation of a functional of the solution of an SDE with some weight, namely the so-called Bismut-Elworthy-Li (BEL) formula as shown in [21] and extended in [40]. In [50] the authors use techniques from Malliavin calculus to prove BEL formula and employ it for the computation of Δ-sensitivities of financial options.

In many applications, it is very natural to expect that the coefficients of a stochastic differential equation (SDE) may depend on properties of the law of the solution, such as dependence on its moments. Here, we want to extend the formula to mean-field type SDEs following the essence of [50] and show that such generalisation is actually non-trivial, requiring more regularity of the solution in the sense of Malliavin. First, we give a relationship between the Malliavin derivative and the spatial derivative of the solution with respect to the initial condition. Already here we see that such generalisation involves an extra factor which is no longer adapted, thus requiring more (Malliavin) regularity on the solution which is not immediate. Fortunately, if b and σ are Lipschitz then the solution is twice Malliavin differentiable, see [14] and hence a formula using Skorokhod integral can be expected. Using such relation one can find the BEL formula in this context. A merely illustrative example is provided to give a better insight on the effect of mean-field SDEs in the BEL formula.

Notations: Denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^d$, $d \geq 1$. Given a Banach space E, denote by $\|\cdot\|_E$ its associated norm. Let $k, p \geq 0$ integers and $\mathbb{D}^{k,p}$ be the space of k times Malliavin differentiable random variables with all p-moments. Denote by D_s, $s \geq 0$ denote the Malliavin derivative as in [90, Chapter 1, Section 1.2.1] and δ its dual operator (Skorokhod integral). Denote by Dom δ the domain of δ (Skorokhod integrable processes). Denote the trace of a matrix $M \in \mathbb{R}^{d\times d}$ by $\mathrm{tr}(M) := \sum_{j=1}^{d} M_{j,j}$ and by M^* its transpose. For a (weakly) differentiable function $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $(x, y) \mapsto f(x, y)$, denote by ∂_1, respectively by ∂_2, (weak) differentiation w.r.t. the first variable $x \in \mathbb{R}^d$, respectively the second variable $y \in \mathbb{R}^d$.

3.2 The (mean-field) Bismut-Elworthy-Li formula

The object of study is a *mean-field* type *stochastic differnetial equation* (SDE) of the form

$$\begin{aligned} dX_t =& b(t, X_t, \rho_t)dt + \sum_{k=1}^{m} \sigma_k(t, X_t, \pi_t)dW_t^k, \quad X_0^x = x \in \mathbb{R}^d, \quad t \in [0, T] \\ \rho_t :=& E[\varphi(X_t)], \quad \pi_t := E[\psi(X_t)] \end{aligned} \tag{3.1}$$

where $T \in \mathbb{R}$, $T > 0$, $b : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma_k : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $k = 1, \ldots, m$, $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, $\psi : \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions and $W = \{W_t, t \in [0, T]\}$ is an m-dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, P)$ equipped with the natural filtration, denoted by $\{\mathcal{F}_t\}_{t\in[0,T]}$.

We will usually consider the solution as a function of x and hence write X_t^x to stress this fact. Otherwise, we will just write X_t. Moreover, we will assume the following conditions as in [25]

(i) the functions $(t, x, y) \mapsto b(t, x, y)$ and $(t, x, y) \mapsto \sigma_k(t, x, y)$, $k = 1, \ldots, m$ are continuously differentiable with bounded Lipschitz derivatives uniformly with respect to $t \in [0, T]$.

(ii) Assume $d \leq m$ and the matrix $(\sigma_1, \ldots, \sigma_m)$ is uniformly elliptic and admits a right pseudo-inverse.

(iii) The functions φ and ψ are continuously differentiable with bounded Lipschitz derivatives.

The following two propositions prepare for the main results of this note.

Proposition 3.1. *Let $X = \{X_t, t \in [0, T]\}$ be the unique global strong solution of (3.1). Then the function $x \mapsto X_t^x$ is continuously differentiable.*

Proof. See [25]. □

Proposition 3.2. *Let $Y = \{Y_t, t \in [0, T]\}$ be the solution to the following matrix-valued linear SDE*

$$dY_t = A_t Y_t dt + \sum_{k=1}^{m} B_t^k Y_t dW_t^k, \quad Y_0 = I, \quad t \in [0, T]$$

where $A_t := \partial_1 b(t, X_t^x, \rho_t^x)$ and $B_t^k := \partial_1 \sigma_k(t, X_t^x, \pi_t^x)$, $k = 1, \ldots, m$. Then

$$(\det Y_t)^{-1} \in \bigcap_{p\geq 1} L^p(\Omega).$$

As a consequence, Y_t is P-a.s. invertible for every $t \in [0, T]$.

Proof. We want to show that

$$E\left[|\det Y_t|^{-p}\right] < \infty$$

for every integer $p \geq 1$. Indeed, in virtue of (stochastic) Liouville's formula which can be found in [106] one has

$$\det Y_t = \exp\left\{\int_0^t \left(\operatorname{tr} A_u + \frac{1}{2}\sum_{k=1}^m (\operatorname{tr} B_u^k)^2\right) du + \sum_{k=1}^m \int_0^t \operatorname{tr} B_u^k dW_u^k\right\}, \quad P-a.s.$$

Hence, by using the property that $E[M_t] = E[M_0] = c$ for a martingale M_t. The claim can be reduced to show that

$$\sup_{t\in[0,T]}\left|E\left[\exp\left\{\lambda\int_0^t \operatorname{tr} A_u du\right\}\right]\right| + \sup_{t\in[0,T]}\left|E\left[\exp\left\{\lambda\int_0^t \sum_{k=1}^m (\operatorname{tr} B_u^k)^2 du\right\}\right]\right| < \infty$$

for every $\lambda \in \mathbb{R}$ which holds since A and B^k, $k = 1, \dots, m$ are uniformly bounded. □

The following is one of the main observations for the derivation of the Bismut-Elworthy-Li formula in the mean-field context.

Theorem 3.3. *Let $X = \{X_t, t \in [0,T]\}$ be the solution of* (3.1). *Then for every $s,t \in [0,T]$, $s \leq t$ one has the following relationship between the spatial derivative and the Malliavin derivative of X_t^x*

$$\frac{\partial}{\partial x}X_t^x = D_s X_t^x \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s \left(I + \int_0^t Y_u^{-1}\left(\alpha_u - \sum_{k=1}^m B_u^k \beta_u^k\right) du + \sum_{k=1}^m \int_0^t Y_u^{-1}\beta_u^k dW_u^k\right) \tag{3.2}$$

for $s \leq t$ where σ^{-1} denotes the right pseudo-inverse of σ, $Y = \{Y_t, t \in [0,T]\}$ is the $d \times d$ fundamental matrix satisfying

$$dY_t = A_t Y_t dt + \sum_{k=1}^m B_t^k Y_t dW_t^k, \quad Y_0 = I, \quad t \in [0,T]$$

and where $A = \{A_t, t \in [0,T]\}$, $\alpha = \{\alpha_t, t \in [0,T]\}$, $B^k = \{B_t^k, t \in [0,T]\}$, $\beta^k = \{\beta_t^k, t \in [0,T]\}$, $k = 1, \dots, m$ are matrix valued processes defined as:

$$\begin{aligned} A_t &:= \partial_1 b(t, X_t^x, \rho_t^x), \quad \alpha_t := \partial_2 b(t, X_t^x, \rho_t^x)\frac{\partial}{\partial x}\rho_t^x \\ B_t^k &:= \partial_1 \sigma_k(t, X_t^x, \pi_t^x), \quad \beta_t^k := \partial_2 \sigma_k(t, X_t^x, \pi_t^x)\frac{\partial}{\partial x}\pi_t^x \end{aligned}$$

for $k = 1, \dots, m$.

Proof. Differentiating with respect to $x \in \mathbb{R}^d$ we have that $\frac{\partial}{\partial x}X_t^x$ satisfies the following matrix-valued linear equation

$$\begin{aligned}\frac{\partial}{\partial x}X_t^x &= I + \int_0^t \left(\partial_1 b(u, X_u^x, \rho_u^x)\frac{\partial}{\partial x}X_u^x + \partial_2 b(u, X_u^x, \rho_u^x)\frac{\partial}{\partial x}\rho_u^x\right) du \\ &\quad + \sum_{k=1}^m \int_0^t \left(\partial_1 \sigma_k(u, X_u^x, \pi_u^x)\frac{\partial}{\partial x}X_u^x + \partial_2 \sigma_k(u, X_u^x, \pi_u^x)\frac{\partial}{\partial x}\pi_u^x\right) dW_u^k.\end{aligned} \tag{3.3}$$

Using the notations in the statement of the theorem, we can solve (3.3) and express $\frac{\partial}{\partial x}X_t$ as

$$\frac{\partial}{\partial x}X_t = Y_t\left(I + \int_0^t Y_u^{-1}\left(\alpha_u - \sum_{k=1}^m B_u^k\beta_u^k\right) du + \sum_{k=1}^m \int_0^t Y_u^{-1}\beta_u^k dB_u^k\right)$$

where $Y = \{Y_t, t \in [0,T]\}$ is the $d \times d$ fundamental matrix satisfying $Y_0 = I$ and

$$dY_t = A_t Y_t dt + \sum_{k=1}^m B_t^k Y_t dW_t^k, \quad t \in [0,T].$$

By the well-known classical relation, see e.g. [90, Chapter 2, Section 2.3.1], it is true that,

$$Y_t = D_s X_t \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s, \quad s \le t$$

where σ^{-1} denotes the right pseudo-inverse of σ and hence the relation follows. □

Remark 3.4. *For the relation $Y_t = D_s X_t \sigma^{-1}(s, X_s, \pi_s) Y_s$, $s \le t$ to hold in the mean-field setting one also needs the property that $x \mapsto X_t^x$ defines a stochastic semiflow. In the mean-field case we point out that the fact that b, σ, φ and ψ are continuously differentiable with bounded derivatives is enough, see e.g. [70].*

It is shown in [14] that SDE (3.1) is twice Malliavin differentiable when the vector field b does not depend on the law of X and one has additive noise. Nevertheless, using the same method one can prove the same result since the dependence on $E[X_t]$ does not bring stochasticity to the equation. In the sense that the Malliavin derivative of X_t for every fixed $t \in [0,T]$ takes the same form as in the usual linear setting.

Henceforth, we will assume the following technical condition for simplicity.

- There exists a bijection $\Lambda_t : \mathbb{R}^d \to \mathbb{R}^d$, for every $t \in [0,T]$, such that the function $b_* : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ defined as

$$\begin{aligned} b_*(t, x, \rho_t, \varphi_t) &:= \partial_t \Lambda_t(\Lambda_t^{-1}(x)) + \partial_x \Lambda_t(\Lambda_t^{-1}(x))[b(t, \Lambda_t^{-1}(x), \rho_t)] \\ &\quad + \frac{1}{2}\partial_x^2 \Lambda_t(\Lambda_t^{-1}(x))\left[\sum_{i=1}^d \sigma(t, \Lambda_t^{-1}(x), \varphi_t)[e_i], \sum_{i=1}^d \sigma(t, \Lambda_t^{-1}(x), \varphi_t)[e_i]\right] \end{aligned}$$

  where $\{e_i\}_{i=1,\dots,d}$ is a basis of $\mathbb{R}^d$, is Lipschitz continuous uniformly with respect to $t \in [0,T]$.

The reason of the above condition is to use Itô's formula on the process $Z_t = \Lambda(X_t)$ so that Z satisfies an SDE with additive noise for which the results from [14] can be applied. Although, it might seem that the class of such processes is small, it covers a wide variety of models which are relevant in applications, such as for instance geometric-type models.

Proposition 3.5. *Let $X = \{X_t, t \in [0,T]\}$ be the unique global strong solution of (3.1). Then we have*

$$X_t \in \bigcap_{p>1} \mathbb{D}^{2,p}(\Omega)$$

for every $t \in [0,T]$.

Proof. See [14]. □

Proposition 3.6. *Let $t \in [0,T]$ and define $F := \int_0^t Y_u^{-1}\left(\alpha_u - \sum_{k=1}^m B_u^k \beta_u^k\right) du$ and $G := \sum_{k=1}^m \int_0^t Y_u^{-1}\beta_u^k dB_u^k$, then $F, G \in$ Dom δ.*

Proof. Since $\mathbb{D}^{1,2} \subset$ Dom δ, see [90, Proposition 1.3.1.], one needs to show that $F, G \in \mathbb{D}^{1,2}$. Since $Y_t = D_s X_t Y_s$, $s \leq t$ and $X_t \in \mathbb{D}^{2,2}$ for every $t \in [0,T]$ we have together with Proposition 3.2 that $Y_t^{-1} \in \mathbb{D}^{1,2}$ for every $t \in [0,T]$. The result follows since the functions α, B^k and β^k are continuously differentiable with bounded derivatives in the first variable. □

Corollary 3.7. *Let*

$$u(s) := \sigma^{-1}(s, X_s^x, \pi_s^x) Y_s \left(I + \int_0^t Y_u^{-1} \left(\alpha_u - \sum_{k=1}^m B_u^k \beta_u^k \right) du + \sum_{k=1}^m \int_0^t Y_u^{-1} \beta_u^k dW_u^k \right)$$

with $0 \leq s \leq t$. Then $u \in$ Dom (δ).

Proof. Indeed, the process u is the product of an adapted process, hence Skorokhod integrable and by Proposition 3.6 a Skorokhod integrable random variable. □

Theorem 3.8 (Bismut-Elworthy-Li formula)**.** *Let $X = \{X_t, t \in [0,T]\}$ be the unique global strong solution of (3.1). Let $\Phi : \mathbb{R} \to \mathbb{R}_+$ be a measurable function such that $\Phi(X_t^x) \in L^2(\Omega)$. Define the function*

$$v(x) := E[\Phi(X_t^x)].$$

Then

$$\frac{\partial}{\partial x} v(x) = E\left[\Phi(X_t^x) \int_0^t a(s) \left[\sigma^{-1}(s, X_s^x, \pi_s^x) u(s)\right]^* \delta W_s\right]^*$$

where $*$ denotes matrix transposition,

$$u(s) := Y_s \left(I + \int_0^t Y_u^{-1} \left(\alpha_u - \sum_{k=1}^m B_u^k \beta_u^k \right) du + \sum_{k=1}^m \int_0^t Y_u^{-1} \beta_u^k dW_u^k \right),$$

here $a : [0,T] \to \mathbb{R}$ is an integrable function such that $\int_0^T a(s) ds = 1$ and functions α, B^k, β^k, $k = 1, \ldots, m$ are defined as in Theorem 3.3.

Proof. Assume first that Φ is infinitely differentiable with compact support. By Theorem 3.3 we have

$$\frac{\partial}{\partial x} X_t = D_s X_t \sigma^{-1}(s, X_s^x, \pi_s^x) u(s), \quad s \leq t, \quad P-a.s.$$

Then multiplying both sides by the function a and integrating over $s \in [0,t]$ we have

$$\frac{\partial}{\partial x} X_t^x = \int_0^t a(s) D_s X_t \sigma^{-1}(s, X_s^x, \pi_s^x) u(s) ds, \quad P-a.s. \tag{3.4}$$

As a consequence,

$$\begin{aligned}
\frac{\partial}{\partial x} v(x) &= E\left[\Phi'(X_t^x) \frac{\partial}{\partial x} X_t^x\right] \\
&= E\left[\Phi'(X_t^x) \int_0^t a(s) D_s X_t \sigma^{-1}(s, X_s^x, \pi_s^x) u(s) ds\right] \\
&= E\left[\int_0^t a(s) D_s \Phi(X_t) \sigma^{-1}(s, X_s^x, \pi_s^x) u(s) ds\right] \\
&= E\left[\Phi(X_t) \int_0^t a(s) \left[\sigma^{-1}(s, X_s^x, \pi_s^x) u(s)\right]^* \delta B_s\right]^*
\end{aligned}$$

where we have used relation (3.4), the chain rule for the Malliavin derivative (backwards) and the duality formula for the Malliavin derivative which is justified by Corollary 3.7.

To extend the formula to bounded functions, one can use a limit argument. For general functions Φ one can use a monotone class argument using the given relation. $\square$

Example 3.9 (Black-Scholes model with continuous dividend payments)**.** *Let $S = \{S_t^x, t \in [0,T]\}$ represent the price dynamics of some asset with initial price $x > 0$ governed by the following SDE*

$$\frac{dS_t^x}{S_t^x} = (\mu - q\rho_t^x)dt + \sigma dB_t, \quad \rho_t^x := E[S_t^x], \quad t \in [0,T], \quad X_0^x = x > 0 \tag{3.5}$$

where $\mu, q, \sigma \in \mathbb{R}$ and $\sigma > 0$. Let $S_t^0 = e^{rt}$, $t \in [0,T]$, $r \in \mathbb{R}$ with $r > 0$ be the risk-less asset and $\Phi : \mathbb{R} \to [0,\infty)$ a pay-off function, then the price of a European option at current time with maturity $T > 0$ (under the risk-neutral valuation approach) is given by

$$p_T(x) = e^{-rT} E_{\widetilde{P}}\left[\Phi(S_T^x)\right]$$

where $\widetilde{P}$ is the risk-neutral measure, i.e.

$$\left.\frac{d\widetilde{P}}{dP}\right|_{\mathcal{F}_t} = M_t^x := e^{-\int_0^t \theta_u^x dW_u - \frac{1}{2}\int_0^t (\theta_u^x)^2 du}, \quad t \in [0,T]$$

where

$$\theta_t^x := \frac{\mu - r - q\rho_t^x}{\sigma}, \quad t \in [0,T].$$

It follows that $\rho_t^x = \frac{x\mu e^{\mu t}}{qxe^{\mu t} + \mu - qx}$ solution to a Riccati equation. Also, $\frac{\partial}{\partial x}\rho_T^x = \frac{\rho_t^x}{x}$ and $\frac{\partial}{\partial x}\theta_t^x = -\frac{q}{x\sigma}\rho_t^x$. Then the Δ-sensitivity of an option Φ on S_T^x is given by

$$\Delta = e^{-rT} E\left[\Phi(S_T^x)\left(\frac{\partial}{\partial x} M_T^x + \frac{1}{x\sigma}\left(1 - q\int_0^T \rho_u^x du\right)\int_0^T a(s) M_T^x \delta W_s\right)\right].$$

Using the integration by parts formula for the Skorokhod integral we find that

$$\int_0^T M_T^x \delta W_s = \left(W(T) + \int_0^T \theta_s^x ds \right) M_T^x$$

and hence, taking $a \equiv \frac{1}{T}$ *we find that under the risk-neutral measure* $\widetilde{P}$, *the* Δ*-sensitivity is given by*

$$\Delta = e^{-rT} E_{\widetilde{P}} \left[\Phi(S_T^x) Z_T \right]$$

with Malliavin weight

$$Z_T := \frac{1}{Tx\sigma} \left(qT \int_0^T \rho_u^x dW_u + qT \int_0^T \theta_u^x \rho_u^x du + \left(1 - q \int_0^T \rho_u^x du \right) \left(W(T) + \int_0^T \theta_s^x ds \right) \right).$$

Finally, observe that if $\rho_t^x \equiv 0$ *then*

$$\Delta = e^{-rT} E_{\widetilde{P}} \left[\Phi(S_T^x) \frac{W(T) + \int_0^T \theta_s^x ds}{Tx\sigma} \right] = e^{-rT} E_P \left[\Phi(S_T^x) \frac{\hat{W}(T)}{Tx\sigma} \right],$$

where $\hat{W}$ *is a standard Brownian motion under* P *and hence the* Δ *coincides with the classical one.*

Chapter 4

Optimal bounds for the densities of solutions of SDEs with measurable and path dependent drift coefficients

David R. Baños and Paul Krühner

Abstract: We consider a process given as the solution of a stochastic differential equation with irregular, path dependent and time-inhomogeneous drift coefficient and additive noise. Explicit and optimal bounds for the Lebesgue density of that process at any given time are derived. The bounds and their optimality is shown by identifying the worst case stochastic differential equation. Then we generalise our findings to a larger class of diffusion coefficients.

4.1 Introduction

The study of regularity of solutions of stochastic differential equations (SDEs) has been a topic of great interest within stochastic analysis, especially since Malliavin calculus was founded. One of the main motivations of Malliavin calculus is precisely to study the regularity properties of the law of Wiener functionals, for instance, solutions to SDEs, as well as, properties of their densities. A classical result on this subject is that if the coefficients of an SDE are C^∞ functions with bounded derivatives of any order and the so-called Hörmander's condition (see e.g. [59]) holds, then the solution of the equation is smooth in the Malliavin sense. Then P. Malliavin shows in [78] that the laws of the solutions at any time are absolutely continuous with respect to the Lebesgue measure and the densities are smooth and bounded. Another approach is attributed to N. Bouleau and F. Hirsch where they show in [23] absolute continuity of the finite dimensional laws of solutions to SDEs based on a stochastic calculus of variations in finite dimensions where they use a limit argument. Also, as a motivation of [23], D. Nualart and M. Zakai [93] found related results on the existence and smoothness of conditional densities of Malliavin differentiable random variables.

It appears to be quite difficult to derive regularity properties for the densities of solutions to SDEs with singular coefficients, i.e. non-Lipschitz coefficients, in particular in the drift. Nevertheless, some findings in this direction have been attained. Let us for instance remark here the work by M. Hayashi, A. Kohatsu-Higa and G. Yûki in [57] where the authors show that SDEs with Hölder continuous drift and smooth elliptic diffusion coefficients admit Hölder continuous densities at any time. Their techniques are mainly based on an integration by parts formula (IPF) in the Malliavin setting and estimates on the characteristic function of the solution in connection with Fourier's inversion theorem. Another result in this direction is due to S. De Marco in [31] where the author proves smoothness of the density on an open domain under the usual condition of ellipticity and that the coefficients are smooth on such domain. A remarkable fact is that Hörmander's condition is skipped in this proof. Moreover, estimates for the tails are also given. The technique relies strongly on Malliavin calculus and an IPF together with estimates on the Fourier transform of the solution. One may already observe that integration by parts formulas in the Malliavin context are a powerful tool for the investigation of densities of random variables as it is the case in the work by V. Bally and L. Caramellino in [5] where an IPF is derived and the integrability of the weight obtained in the formula gives the desired regularity of the density. As a consequence of the aforesaid result D. Baños and T. Nilssen give in [14] a criterion to obtain regularity of densities of solutions to SDEs according to how regular the drift is. The technique is also based on Malliavin calculus and a sharp estimate on the moments of the derivative of the flow associated to the solution. This result is a slight improvement of a very similar criterion obtained by S. Kusuoka and D. Stroock in [73] when the diffusion coefficient is constant and the drift may be unbounded. Another related result on upper and lower bounds for densities is due to V. Bally and A. Kohatsu-Higa in [6] where bounds for the density of a type of a two-dimensional degenerated SDE are obtained. For this case, it is assumed that the coefficients are five times differentiable with bounded derivatives. Finally, we also mention the results by A. Kohatsu-Higa and A. Makhlouf in [64] where the authors show smoothness of the density for smooth coefficients that may also depend on an external process whose drift coefficient is irregular. They also give upper and lower estimates for the density.

It is worth alluding the exceptional result by A. Debussche and N. Fournier in [32] on this topic where the authors show that the finite dimensional densities of a solution of an SDE with jumps lies in a certain (low regular) Besov space when the drift is Hölder continuous. The novelty is that their method does not use Malliavin calculus as in the aforementioned works.

It is therefore important to highlight that in this paper we do *not* use Malliavin calculus or any other type of variational calculus and we see this as an alternative perspective for studying similar problems. Instead, we employ control theory techniques to, shortly speaking, reduce the overall problem to a critical case for which many results in the literature are available. In particular, our technique entitles us to find a *worst case* SDE whose solution has an explicit density that dominates all densities of solutions to SDEs among those with measurable bounded drifts.

We believe this method is robust since no well-behaviour on the drift is needed other than merely boundedness and no Markovianity of the system is assumed. Certainly, no regularity is obtained but we are confident that the method can be exploited to gain more regularity of the densities.

This paper is organised as follows. In Section 4.2 we summarise our main results with some generalisations to non-trivial diffusion coefficients and to any arbitrary dimension. We also give some insight on concrete properties of the bounds as well as some examples with graphics. Section 4.3 is devoted to thoroughly prove the assertions of the main results. More specifically, we will give an argument based on a control problem to reduce the problem to one critical case. We will also prove in detail the properties adduced in the previous section.

4.1.1 Notations

We denote the strictly positive numbers by $\mathbb{R}_{++} := (0, \infty)$, the trace of a matrix $M \in \mathbb{R}^{d\times d}$ by $\mathrm{tr}(M) := \sum_{j=1}^{d} M_{j,j}$ and $\pm$ simply denotes either $+$ or $-$. The Skorokhod space $\mathbb{D}(\mathbb{R}^d)$ is the set of all càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}^d$ equipped with the Skorokhod metric, c.f. [60, Chapter VI.1]. The canonical space is the triplet $(\mathbb{D}(\mathbb{R}^d), \mathfrak{D}, (\mathcal{D}_t)_{t\geq 0})$ where $\mathfrak{D}$ is the σ-algebra generated by the point evaluations and $(\mathcal{D}_t)_{t\geq 0}$ is the right-continuous filtration generated by the canonical process $X : \mathbb{R}_+ \times \mathbb{D}(\mathbb{R}^d) \to \mathbb{R}^d, (t, f) \mapsto f(t)$. We denote the generalised signum function by $\mathrm{sgn}(x) := 1_{\{x\neq 0\}} x/|x|$ for any $x \in \mathbb{R}^d$. This is the orthogonal projection to the unit Euclidean sphere. For a complex number $z \in \mathbb{C}$ we denote its real resp. imaginary part by $\mathrm{Re}(z)$ resp. $\mathrm{Im}(z)$.

Further notations are used as in [60].

4.2 Main results

In this section we present our main result and some direct consequences. In particular, we will find sharp explicit bounds for SDEs with additive noise in the one-dimensional case and give some extensions to the d-dimensional case with more general diffusion coefficients.

Throughout this section let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathfrak{A}, P)$ be a filtered probability space with the usual assumptions on the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$, i.e. $\mathcal{F}_0$ contains all P-null sets and $\mathcal{F}$ is right-continuous, W be a d-dimensional standard Brownian motion and we define the process classes

$$\begin{aligned} \mathcal{A}_+ &:= \{u : u \text{ is a stochastic process bounded by } 1\} \\ \mathcal{A} &:= \{u \in \mathcal{A}_+ : u \text{ is } \mathcal{F}\text{-adapted}\}. \end{aligned}$$

The next results constitutes one of the core results of this section and will be proven in detail in the next section.

Theorem 4.1. *Let $C > 0$, W be a d-dimensional standard Brownian motion and $u \in \mathcal{A}$. Then $X(t) := \int_0^t Cu(s)ds + W(t)$ has Lebesgue density*

$$\rho_t(x) := \limsup_{\epsilon\to 0} \frac{P(|X(t) - x| \leq \epsilon)}{V_\epsilon},\ x \in \mathbb{R}^d$$

where $V_\epsilon = \frac{\pi^{d/2}}{\Gamma(d/2+1)}\epsilon^d$ denotes the volume of the d-dimensional Euclidean ball with radius ϵ and Γ denotes the gamma function. Moreover, ρ_t satisfies

$$0 < \alpha_{d,t,C}(x) \leq \rho_t(x) \leq \beta_{d,t,C}(x) \leq \beta_{d,t,C}(0)$$

for any $t > 0$, $x \in \mathbb{R}$ where

$$\alpha_{d,t,C}(x) := \limsup_{\epsilon\to 0} \frac{P(|Y^+_{Cx}(tC^2)| \leq C\epsilon)}{V_\epsilon},\quad \beta_{d,t,C}(x) := \limsup_{\epsilon\to 0} \frac{P(|Y^-_{Cx}(tC^2)| \leq C\epsilon)}{V_\epsilon},$$

and Y_x^+ and Y_x^- are the unique solutions to the SDEs

$$Y_x^+(t) = x + \int_0^t \operatorname{sgn}(Y_x^+(s))ds + W(t),$$
$$Y_x^-(t) = x - \int_0^t \operatorname{sgn}(Y_x^-(s))ds + W(t)$$

for any $t \geq 0$.

Proof. See at the end of Section 4.3. □

If $d = 1$, then the functions α, β as well as some of their properties can be derived explicitly, cf. Theorem 4.10. In the multidimensional case we can give some of their properties. Let us summarise the formulas.

Theorem 4.2. *Let $t > 0$, $C > 0$ and α, β be given as in Theorem 4.1. Then*

$$\alpha_{1,t,C}(0) = \frac{1}{\sqrt{t}}\varphi\left(C\sqrt{t}\right) - C\Phi\left(-C\sqrt{t}\right), \quad \text{and}$$
$$\beta_{1,t,C}(0) = \frac{1}{\sqrt{t}}\varphi\left(C\sqrt{t}\right) + C\Phi\left(C\sqrt{t}\right)$$

where Φ resp. φ denotes the distribution resp. density function of the standard normal law. For $x \in \mathbb{R}\backslash\{0\}$ we have

$$\alpha_{1,t,C}(x) = \int_0^{tC^2} C\alpha_{1,tC^2-s,1}(0)\rho_{\theta_0^{Cx}}(s)ds \quad \text{and}$$
$$\beta_{1,t,C}(x) = \int_0^{tC^2} C\beta_{1,tC^2-s,1}(0)\rho_{\tau_0^{Cx}}(s)ds$$

where

$$\rho_{\tau_0^x}(t) = \frac{|x|}{\sqrt{2\pi t^3}}e^{-\frac{(|x|-t)^2}{2t}} \quad \text{and}$$
$$\rho_{\theta_0^x}(t) = \frac{|x|}{\sqrt{2\pi t^3}}e^{-\frac{(|x|+t)^2}{2t}}$$

for any $s > 0$. Moreover, we have

$$\frac{2^d}{C_d d^{d/2}}\prod_{i=1}^d \alpha_{1,t,C}(x_i) \leq \alpha_{d,t,C}(x) \leq \beta_{d,t,C}(x) \leq \frac{2^d}{C_d}\prod_{i=1}^d \beta_{1,t,C}(x_i), \quad x \in \mathbb{R}^d$$

where $C_d := \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$ for any $x \in \mathbb{R}^d$.

Proof. This is part of the statements of Theorems 4.10 and 4.12 below. □

In what follows, we will derive bounds for the densities of solutions of general SDEs. The following is an immediate consequence of Theorem 4.1.

Corollary 4.3. *Let $C > 0$, $x_0 \in \mathbb{R}^d$, $b : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}$ be predictable and bounded by C. Then any weak solution of the SDE*

$$X(t) = x_0 + \int_0^t b(s, X)ds + W(t), \quad t \geq 0$$

has density ρ_t at time $t > 0$ which is bounded from below by $x \mapsto \alpha_{d,t,C}(x - x_0)$ and from above by $x \mapsto \beta_{d,t,C}(x - x_0)$ where α and β are given in Theorem 4.1 and W is a d-dimensional Brownian motion. Moreover, the bounds are optimal in the sense that for any $x_1, x_2 \in \mathbb{R}^d$ there are two functionals b_{x_1}, resp. b_{x_2} for which the density ρ_t of the solution to the SDE $dX(t) = b_{x_1}(X(t))dt + W(t), X(0) = 0$, resp. $dX(t) = b_{x_2}(X(t))dt + W(t), X(0) = 0$ attains the upper bound in x_1, resp. the lower bound in x_2.

Proof. Define $Y(t) := X(t) - x_0$ and $u(t) := b(t, X)$ for any $t \geq 0$. Then

$$Y(t) = \int_0^t u(s)ds + W(t), \quad t \geq 0.$$

The bounds follow from Theorem 4.1. Shifts of the processes Y^-, resp. Y^+ attain the upper, resp. lower bounds at the given points. □

Now we focus on our second main result which is an application of Corollary 4.3. This time X is given as a solution of an SDE with measurable drift and a diffusion coefficient which is continuously differentiable.

Theorem 4.4. *Let $b : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^d$ be predictable, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$ be continuously differentiable and assume the following conditions.*

1. *$\sigma(t, x)$ is an invertible matrix for any $t \geq 0$, $x \in \mathbb{R}^d$.*

2. *There is a function $F : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ such that $D_2F(t, x) = (\sigma(t, x))^{-1}$ for any $t \geq 0$, $x \in \mathbb{R}^d$ where $D_2F(t, x)$ denotes the Fréchet derivative of $F(t, \cdot)$ with respect to x.*

3. *The function*

$$\begin{aligned} \widetilde{b} &: \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^d, \\ (t, f) &\mapsto \partial_1 F(t, f(t)) + \sigma(t, f(t))^{-1} b(t, f) \\ &\qquad + \frac{1}{2} \Big(\mathrm{tr}\, \big(\, \sigma(t, f(t))^\top H_2 F_k(t, f(t)) \sigma(t, f(t)) \,\big)\Big)_{k=1,\dots,d} \end{aligned}$$

   is bounded by some constant $C > 0$ where $H_2F_k(t, x)$ denotes the Hessian matrix of $F_k(t, \cdot)$, i.e. $(\partial_{x_i}\partial_{x_j} F_k(t, x))_{i,j=1,\dots,d}$ for any $t \geq 0$, $x \in \mathbb{R}^d$.

Then any solution of the SDE

$$X(t) = x_0 + \int_0^t b(s, X)ds + \int_0^t \sigma(s, X(s))dW(s)$$

has, at each time t, Lebegsue density ρ_t and for every $x \in \mathbb{R}^d$ we have

$$\rho_t(x) \leq \frac{\beta_{d,t,C}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))}$$

where $\alpha_{d,t,C}$, $\beta_{d,t,C}$ are defined as in Theorem 4.1. Moreover, if additionally $F(t,\cdot)$ is invertible for any fixed $t > 0$, then

$$0 < \frac{\alpha_{d,t,C}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))} \leq \rho_t(x) \leq \frac{\beta_{d,t,C}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))}.$$

Proof. Define $Y(t) := F(t, X(t))$ and $u(t) := \widetilde{b}(t, X)$ for any $t \geq 0$. Then Itô's formula yields

$$Y(t) = F(0,x_0) + \int_0^t u(s)ds + W(t), \quad t \geq 0.$$

Theorem 4.1 states that $Y(t)$ has Lebesgue density $\rho_{Y(t)}$ which admits the bounds

$$\alpha_{d,t,C}(y - F(0,x_0)) \leq \rho_{Y(t)}(y) \leq \beta_{d,t,C}(y - F(0,x_0))$$

for any $t > 0$, $y \in \mathbb{R}^d$.

From the definition of $Y(t)$ we directly get

$$\rho_t(x) \leq \frac{\rho_{Y(t)}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))} \leq \frac{\beta_{d,t,C}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))}$$

for any $t > 0$, $x \in \mathbb{R}^d$.

If we assume that $F(t,\cdot)$ is invertible for any $t > 0$, then

$$\rho_t(x) = \frac{\rho_{Y(t)}(F(t,x) - F(0,x_0))}{\operatorname{tr}(\sigma(t,x))}$$

for any $x \in \mathbb{R}^d$ and, hence, the additional claim follows. □

The conditions (1) to (3) appearing in Theorem 4.4 simplify considerably in dimension 1. Moreover, due to Itô-Tanaka's formula we can relax the conditions on σ.

Theorem 4.5. *Let X be a solution of the SDE*

$$X(t) = x_0 + \int_0^t b(s,X)dt + \int_0^t \sigma(X(s))dW(s)$$

where $x_0 \in \mathbb{R}$, W is a standard Brownian motion, $b : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$ predictable and bounded by some constant C_b, $\sigma : \mathbb{R} \to \mathbb{R}_+$ is a Lipschitz continuous function with Lipschitz bound L and $\sigma(x) \geq \epsilon$ for some constant $\epsilon > 0$.

Then $X(t)$ has Lebesgue density ρ_t and

$$0 < \frac{\alpha_{t,C}(|F(x) - F(x_0)|)}{\sigma(x)} \leq \rho_t(x) \leq \frac{\beta_{t,C}(|F(x) - F(x_0)|)}{\sigma(x)}$$

for any $t > 0$ where $\alpha_{t,C}$ and $\beta_{t,C}$ are defined as in Theorem 4.1 when $d = 1$, $F(x) := \int_0^x \frac{1}{\sigma(u)} du$ and

$$C := \sup \left\{ \left| \frac{b(t,f)}{\sigma(f(t))} \right| : t \in \mathbb{R}_+, f \in C(\mathbb{R}_+, \mathbb{R}) \right\} + L/2.$$

Moreover, $C \leq \frac{C_b}{\epsilon} + L/2$ where C_b is a uniform bound for b.

Proof. Define $Y(t) := F(X(t))$. Since σ is Lipschitz continuous there is a function $\sigma' : \mathbb{R}_+ \to \mathbb{R}$ which is bounded by L and $\sigma(x) = \sigma(0) + \int_0^x \sigma'(u)du$. Then Itô-Tanaka's formula [100, Theorem VI.1.5] yields

$$Y(t) = F(x_0) + \int_0^t \left(\frac{b(s,X)}{\sigma(X(s))} - \frac{1}{2}\sigma'(X(s)) \right) ds + W(t).$$

Let $G := F^{-1}$ and define

$$\widetilde{b}(s,y) := \frac{b(s, G \circ f)}{\sigma(G(f(s))} - \frac{1}{2}\sigma'(G(f(s))), \quad s \in \mathbb{R}_+, f \in C(\mathbb{R}_+, \mathbb{R})$$

which is predictable and bounded by C. Then the result follows from Corollary 4.3. $\square$

In the next section we will give precise definitions and mathematical computations of the functions $\alpha_{d,t,C}$ and $\beta_{d,t,C}$ in dimension 1 and why these are the optimal bounds (in the sense of Corollary 4.3) for the densities of SDEs with bounded measurable drifts. Before we do that, let us give some intuitive insight on the shape and behaviour of these bounds for the one-dimensional case. Consider any one-dimensional process of the form

$$X(t) = \int_0^t u(s)ds + W(t), \quad t \geq 0, \quad u \in \mathcal{A}$$

as in Theorem 4.1. In particular, X can be the solution to the following SDE, $dX(t) = b(t,X)dt + dW(t)$, $X(0) = 0$, $t \geq 0$, with b bounded and predictable as in Corollary 4.3. Furthermore, denote by ρ_t the density of $X(t)$ at a fixed time $t > 0$. Then Theorem 4.1 grants that $0 < \alpha_t(x) \leq \rho_t(x) \leq \beta_t(x)$ for any $x \in \mathbb{R}$. In the following figure we can observe the functions α_t and β_t for different values of $t > 0$ and see how they behave. We can see the function α_t in orange and β_t in green. Any density lies between these two curves and these bounds are optimal in the sense that, for given $x_0, y_0 \in \mathbb{R}$ we can find drifts u_{x_0} and u_{y_0} such that the associated densities $\rho_t^{x_0}$, resp. $\rho_t^{y_0}$ for these drift coefficients satisfy $\rho_t(x_0) = \alpha_t(x_0)$, respectively, $\rho_t(y_0) = \beta_t(y_0)$. As an illustration we just take the drift to be $+\mathrm{sgn}(x - 0.25)$ in blue and $-\mathrm{sgn}(x - 1)$ in red.

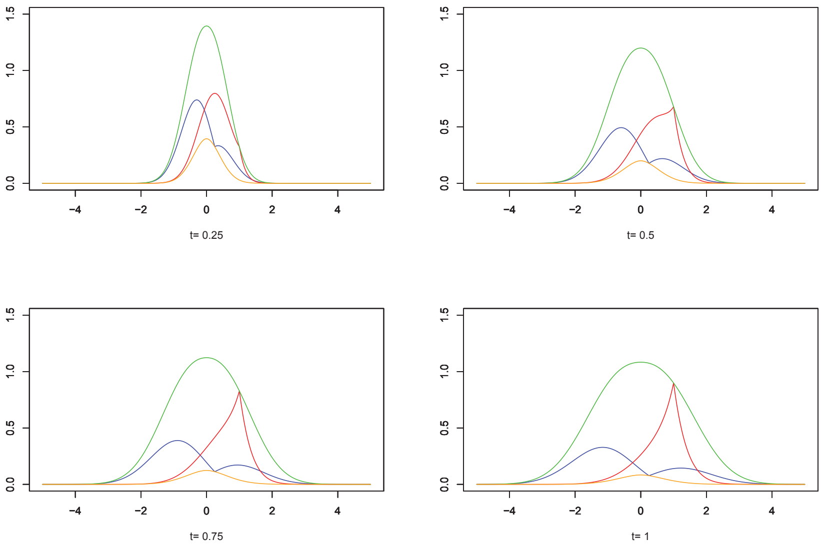

Figure 4.1: Upper and lower bounds for $C = 1$ starting at $x = 0$ (in green and orange) with the respective densities when the drift coefficients are $\mathrm{sgn}(x - 0.25)$ and $-\mathrm{sgn}(x - 1)$ (blue and red) at different times $t \in \{0.25, 0.5, 0.75, 1\}$.

As we can see, both densities are bounded by α_t and β_t and the bounds are attained in 0.25 for density of the process with drift $+\mathrm{sgn}(x - 0.25)$ (in blue) and in 1 when the drift is $-\mathrm{sgn}(x - 1)$ (in red).

4.3 Reduction and the critical case

In this section we will see how to derive the functions $\alpha_{t,C}$ and $\beta_{t,C}$ explicitly for the case $d = 1$ as well as some of their properties, cf. Theorem 4.10. Then we will show that these are indeed the bounds for the densities of any solution to SDEs with bounded measurable drift by solving a stochastic control problem, cf. Theorem 4.18 and thereafter we give the proof for Theorem 4.1. In the sequel, consider the process

$$Y_x^{\pm}(t) := x \pm \int_0^t \mathrm{sgn}(Y_x^{\pm}(s))ds + W(t), \quad t \geq 0, \tag{4.1}$$

c.f. [105] for existence and (pathwise) uniqueness. Moreover, at some point we will also use the property that the solution to equation (4.1) is strong Markov, even for the multidimensional case. This can be for instance justified using [3, Theorem 6.4.5] in connection with [100, Corollary IX.1.14].

Lemma 4.6. *For every $t > 0$, $Y_0^+(t)$ resp. $Y_0^-(t)$ has density $\rho_{Y_0^+(t)}$, resp. $\rho_{Y_0^-(t)}$ given by*

$$p_t(0,y) := \rho_{Y_0^+(t)} = \frac{1}{\sqrt{t}}\varphi\left(\frac{|y|-t}{\sqrt{t}}\right) - e^{2|y|}\Phi\left(-\frac{|y|+t}{\sqrt{t}}\right), \quad \textit{resp.}$$
$$q_t(0,y) := \rho_{Y_0^-(t)} = \frac{1}{\sqrt{t}}\varphi\left(\frac{t+|y|}{\sqrt{t}}\right) + e^{-2|y|}\Phi\left(\frac{t-|y|}{\sqrt{t}}\right)$$

for $y \in \mathbb{R}$ and any $t > 0$ where φ, resp. Φ, denote the density, resp. the distribution function, of the standard normal law.

Proof. The density for $Y_0^-(t)$ is the statement of [62, Exercise 6.3.5] as for $Y_0^+(t)$ computations are fairly similar. $\square$

The computation of the densities $\rho_{Y_0^+(t)}$ and $\rho_{Y_0^-(t)}$ in the previous lemma are relatively easy given the fact that the local-time of the Brownian motion starting from 0 is symmetric and the joint law of $W(t)$ and the local time of W, $L_t^W(0)$ is explicitly known, see [62]. Nevertheless, one is able to find reasonably explicit expressions for the densities of $Y_x^+(t)$ and $Y_x^-(t)$ which yield representations for α and β if $d = 1$.

First we focus on the computation of the density of $Y_x^-(t)$ and then for $Y_x^+(t)$ which is similar.

Lemma 4.7. *For every $t \geq 0$, the density of $Y_x^-(t)$ is given by*

$$q_t(x,y) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{(\mathrm{sgn}(x)(x-y)-t)^2}{2t}}\left(1 - e^{-\frac{2xy}{t}}\right)1_{\{\mathrm{sgn}(xy)\geq 0\}} + \int_0^t q_{t-s}(0,y)\rho_{\tau_0^x}(s)ds$$

where $x, y \in \mathbb{R}$, $x \neq 0$ and τ_0^x is the first hitting time of the process $Y_x^-(t)$ at 0 whose density function is explicitly given by

$$\rho_{\tau_0^x}(s) = \frac{|x|}{\sqrt{2\pi s^3}}e^{-\frac{(|x|-s)^2}{2s}}, \; s > 0.$$

Proof. Let τ_0^x be the first time the process Y_x^- hits 0, i.e.

$$\tau_0^x := \inf\{t \geq 0 : \; Y_x^-(t) = 0\}.$$

Then it is clear, that $Y_x^-(t) = x - \mathrm{sgn}(x)t + W(t)$ for any $t \in [0, \tau_0^x]$. Define $\widetilde{W} := -W$ and $B(t) := \mathrm{sgn}(x)t + \widetilde{W}(t)$. The process $B(t)$ is a Brownian motion with drift starting at 0. It is clear, that $\tau_0^x = \inf\{t \geq 0 : B(t) = x\}$, whose law is known, namely τ_0^x is inverse Gaussian distributed and [22, p.223, Formula 2.0.2] states that its density is given by

$$\rho_{\tau_0^x}(t) = \frac{|x|}{\sqrt{2\pi t^3}}e^{-\frac{(|x|-t)^2}{2t}}, \; t > 0.$$

Now define $f_\varepsilon(z) := \frac{1}{2\varepsilon}1_{(y-\varepsilon,y+\varepsilon)}(z)$ for a fixed $y \in \mathbb{R}$, then

$$\begin{aligned} \mathrm{E}[f_\varepsilon(Y_x(t))] &= \mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t<\tau_0^x\}}] + \mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t\geq\tau_0^x\}}] \\ &= A_1 + A_2 \end{aligned}$$

where $A_1 := \mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t<\tau_0^x\}}]$ and $A_2 := \mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t\geq\tau_0^x\}}]$. We have

$$\begin{aligned} P\left(Y_x^-(t)\leq y, t<\tau_0^x\right) &= P\left(x-\operatorname{sgn}(x)t+W(t)\leq y, t<\tau_0^x\right) \\ &= P\left(B(t)\geq x-y, t<\tau_0^x\right). \end{aligned}$$

We start with the case $x>0$. Observe that $\tau_0^x = \inf\{t>0 : B(t)=x\}$ and hence $\{t<\tau_0^x\} = \{M(t)<x\}$ where $M(t) := \sup_{s\in[0,t]} B(s)$. As a consequence

$$\begin{aligned} P\left(Y_x^-(t)\leq y, t<\tau_0^x\right) &= P\left(B(t)\geq x-y, M(t)<x\right) \\ &= \mathrm{E}\left[1_{\{B(t)\geq x-y, M(t)<x\}}\right] \\ &= \mathrm{E}_Q\left[1_{\{B(t)\geq x-y, M(t)<x\}}\frac{1}{Z(t)}\right] \end{aligned}$$

where Q is the equivalent measure w.r.t. P defined by

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \exp\left\{-\operatorname{sgn}(x)\widetilde{W}(t)-t/2\right\} =: Z(t), \quad t\geq 0.$$

[95, Theorem 8.6.4] yields that the process $B(t) = \operatorname{sgn}(x)t+\widetilde{W}(t)$, $t\geq 0$ is a standard Q-Brownian motion and $M(t)$ is therefore the running maximum of the standard Brownian motion B, hence

$$P\left(Y_x^-(t)\leq y, t\leq\tau_0^x\right) = \int_0^\infty\int_{-\infty}^w 1_{\{z\geq x-y, w<x\}}e^{\operatorname{sgn}(x)z-t/2}\rho_{B(t),M(t)}(z,w)dzdw \tag{4.2}$$

where $\rho_{B(t),M(t)}$ denotes the joint density of $B(t)$ and $M(t)$ which is explicitly given, see [62, Proposition 2.8.1], by

$$\rho_{B(t),M(t)}(z,w) = \frac{2(2w-z)}{\sqrt{2\pi t^3}}\exp\left\{-\frac{(2w-z)^2}{2t}\right\},\ z\leq w,\ w\geq 0.$$

We have

$$\begin{aligned} A_1 &= \frac{1}{2\varepsilon}P\left(y-\varepsilon\leq Y_x^-(t)\leq y+\varepsilon, t\leq\tau_0^x\right) \\ &= \frac{1}{2\varepsilon}\int_0^\infty\int_{-\infty}^w 1_{\{x-y-\varepsilon\leq z\leq x-y+\varepsilon, w<x\}}e^{\operatorname{sgn}(x)z-t/2}\rho_{B(t),M(t)}(z,w)dzdw \end{aligned}$$

Finally, the above probability converges to the derivative of (4.2) w.r.t. y, that is

$$\begin{aligned} &\lim_{\varepsilon\searrow 0}\frac{1}{2\varepsilon}P\left(y-\varepsilon\leq Y_x^-(t)\leq y+\varepsilon, t<\tau_0^x\right) \\ &= e^{\operatorname{sgn}(x)(x-y)-t/2}\int_{x-y}^x\rho_{B(t),M(t)}(x-y,w)dw \\ &= \frac{1}{\sqrt{2\pi t}}e^{\operatorname{sgn}(x)(x-y)-t/2}\left(e^{-(x-y)^2/2t}-e^{-(x+y)^2/2t}\right)1_{\{x\geq x-y\}} \\ &= \frac{1}{\sqrt{2\pi t}}e^{-\frac{(\operatorname{sgn}(x)(x-y)-t)^2}{2t}}\left(1-e^{-\frac{2xy}{t}}\right)1_{\{y\geq 0\}}. \end{aligned}$$

Now we continue to compute A_2. Define the random variable $\tau := \tau_0^x \vee t$. It is readily checked that $\tau \geq \tau_0^x$ and τ is $\mathcal{F}_{\tau_0^x}$-measurable because the event $\{t \geq \tau_0^x\}$ is in $\mathcal{F}_{\tau_0^x}$. Then the strong Markov property of Y_x^- and [62, Corollary 2.6.18] yield

$$\begin{aligned}
\mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t\geq\tau_0^x\}}|\mathcal{F}_{\tau_0^x}] &= \mathrm{E}[f_\varepsilon(Y_x^-(\tau))1_{\{t\geq\tau_0^x\}}|\mathcal{F}_{\tau_0^x}] \\
&= 1_{\{t\geq\tau_0^x\}}\mathrm{E}[f_\varepsilon(Y_x^-(\tau))|\mathcal{F}_{\tau_0^x}] \\
&= 1_{\{t\geq\tau_0^x\}}\mathrm{E}[f_\varepsilon(Y_0^-(\xi))]|_{\xi=\tau-\tau_0^x}
\end{aligned}$$

P-a.s. As a consequence

$$\begin{aligned}
\mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t\geq\tau_0^x\}}] &= \mathrm{E}\left[\mathrm{E}[f_\varepsilon(Y_x^-(t))1_{\{t\geq\tau_0^x\}}|\mathcal{F}_{\tau_0^x}]\right] \\
&= \mathrm{E}\left[1_{\{t\geq\tau_0^x\}}\mathrm{E}[f_\varepsilon(Y_0^-(\xi))]|_{\xi=\tau-\tau_0^x}\right] \\
&= \mathrm{E}\left[1_{\{t\geq\tau_0^x\}}\mathrm{E}[f_\varepsilon(Y_0^-(\xi))]|_{\xi=t-\tau_0^x}\right].
\end{aligned}$$

Now, the density of $Y_0^-(t)$ is explicitly known by Lemma 4.6. Thus

$$A_2 = \mathrm{E}\left[\int_{\mathbb{R}} f_\varepsilon(z)q_{t-\tau_0^x}(0,z)1_{\{t\geq\tau_0^x\}}\right] = \int_0^t\int_{\mathbb{R}} f_\varepsilon(z)q_{t-s}(0,z)\rho_{\tau_0^x}(s)ds.$$

Then, letting $\varepsilon \to 0$ and by Lebesgue's dominated convergence theorem we obtain that, for $x > 0$ and $y \in \mathbb{R}$

$$q_t(x,y) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{(\mathrm{sgn}(x)(x-y)-t)^2}{2t}}\left(1-e^{-\frac{2xy}{t}}\right)1_{\{y\geq 0\}} + \int_0^t q_{t-s}(0,y)\rho_{\tau_0^x}(s)ds.$$

We have

$$\begin{aligned}
-Y_{-x}^-(t) &= x + \int_0^t \mathrm{sgn}(Y_{-x}^-(s))ds + \widetilde{W}(t) \\
&= x - \int_0^t \mathrm{sgn}(-Y_{-x}^-(s))ds + \widetilde{W}(t)
\end{aligned}$$

for any $t \geq 0$ and hence $(-Y_{-x}^-, \widetilde{W})$ is a weak solution of (4.1) for $\pm = -$ and starting point x. Hence, $-Y_{-x}^-(t)$ has the same law as $Y_x^-(t)$ for any $t \geq 0$. Consequently, we have

$$q_t(x,y) = q_t(-y,-x), \quad x > 0, y \in \mathbb{R}.$$

The claimed formula follows. □

Similarly, we can also obtain the density for $Y_x^+(t)$. The proof follows exactly the same ideas as in Lemma 4.7 and has therefore been omitted.

Lemma 4.8. *For every $t \geq 0$, the density of $Y_x^+(t)$ is given by*

$$p_t(x,y) := \frac{2}{\sqrt{2\pi t}}e^{-\frac{(\mathrm{sgn}(x)(x-y)+t)^2}{2t}}\left(1-e^{\frac{-2xy}{t}}\right)1_{\{\mathrm{sgn}(xy)\geq 0\}} + \int_0^t p_{t-s}(0,y)\rho_{\theta_0^x}(s)ds.$$

for $x, y \in \mathbb{R}$, $x \neq 0$ and θ_0^x is the first hitting time of where

$$\rho_{\theta_0^x}(s) = \frac{|x|}{\sqrt{2\pi s^3}} e^{-\frac{(|x|+s)^2}{2s}}, \quad 0 < s < \infty.$$

Proof. The proof of this lemma follows completely the same ideas as in Lemma 4.7. One of the main differences is that in this case the distribution of the stopping time θ_0^x has an atom at infinity, namely, from [22, p.223, Formula 2.0.2] we have

$$\rho_{\theta_0^x}(t) = \frac{|x|}{\sqrt{2\pi t^3}} e^{-\frac{(|x|+t)^2}{2t}}, \ 0 < t < \infty$$

and

$$P(\theta_0^x = \infty) = 1 - e^{-2|x|}.$$

□

Now we are in a position to define the functions $\alpha_{t,C}$ and $\beta_{t,C}$ for the one-dimensional case and study some of their properties. Before we do that, we will need a technical result to prove one of the properties of these functions.

Proposition 4.9. *Let $b : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be bounded and measurable and*

$$X_x(t) := x + \int_0^t b(s, X_x(s))ds + W(t), \quad x \in \mathbb{R}, \quad t \geq 0$$

where W is a 1-dimensional Brownian motion. Then

$$X_x(t) \leq X_y(t) \quad P\text{-a.s.}$$

for any $t \geq 0$, $x, y \in \mathbb{R}$ with $x \leq y$.

Proof. Define

$$Y_x(t) := X_x(t) - W(t) = x + \int_0^t b(s, Y_x(s) + W(s))ds = x + \int_0^t \widetilde{b}(s, Y_x(s))ds$$

where the equalities hold P-a.s. and here $\widetilde{b}(t, z) := b(t, z + W(t))$ for any $t \geq 0$, $z \in \mathbb{R}$. Let $x, y \in \mathbb{R}$ with $x \leq y$ and define $Z(t) := \min\{Y_x(t), Y_y(t)\}$. Then

$$Z(t) = x + \int_0^t \widetilde{b}(s, Z(s))ds, \quad t \geq 0.$$

Hence $U(t) := Z(t) + W(t) = x + \int_0^t b(s, U(s))ds + W(t)$. [100, Theorem IX.3.5 i)] yields $U(t) = X_x(t)$ a.s. Observe that $U(t) = \min\{X_x(t), X_y(t)\}$ and hence

$$X_x(t) = U(t) \leq X_y(t), \quad t \geq 0$$

P-a.s. □

Theorem 4.10. *Let q be the transition density of the Markov process Y^- which is given in Lemma 4.7 and p the transition density for the Markov process Y^+ given in Lemma 4.8. Define the functions $\alpha, \beta : \mathbb{R}_{++} \times \mathbb{R}_+ \times \mathbb{R} \to (0, \infty)$ by $\alpha_{t,C}(x) := Cp_{tC^2}(Cx, 0)$ and $\beta_{t,C}(x) := Cq_{tC^2}(Cx, 0)$ where $t > 0$, $C > 0$ and $x \in \mathbb{R}$. Then*

$$\begin{aligned}\alpha_{t,C}(x) &= \int_0^{tC^2} Cp_{tC^2-s}(0,0)\rho_{\theta_0^{Cx}}(s)ds, \\ &= \int_0^{tC^2} \left(\frac{C}{\sqrt{tC^2-s}}\varphi(\sqrt{tC^2-s}) - C\Phi(-\sqrt{tC^2-s})\right)\rho_{\theta_0^{Cx}}(s)ds,\ x \neq 0,\end{aligned} \tag{4.3}$$

and

$$\begin{aligned}\beta_{t,C}(x) &= \int_0^{tC^2} Cq_{tC^2-s}(0,0)\rho_{\tau_0^{Cx}}(s)ds \\ &= \int_0^{tC^2} \left(\frac{C}{\sqrt{tC^2-s}}\varphi(\sqrt{tC^2-s}) + C\Phi(\sqrt{tC^2-s})\right)\rho_{\tau_0^{Cx}}(s)ds, \quad x \neq 0\end{aligned} \tag{4.4}$$

where recall that $\rho_{\theta_0^x}$, respectively $\rho_{\tau_0^x}$ are given as in Lemma 4.8, respectively as in Lemma 4.7.

In addition, for each $t > 0$ and $C > 0$ the functions $\alpha_{t,C}$ and $\beta_{t,C}$ are analytic in $\mathbb{R} \setminus \{0\}$, Lipschitz continuous in $\mathbb{R}$, symmetric, decreasing on $[0, \infty)$ and by symmetry increasing on $(-\infty, 0]$. They have exponential decay of the type $o(c_1|x|e^{c_2|x|}e^{-c_3|x|^2})$ for constants $c_1, c_2, c_3 > 0$. Moreover, they attain their maxima at $x = 0$ which are given by

$$\alpha_{t,C}(0) = Cp_{tC^2}(0,0) = \frac{1}{\sqrt{t}}\varphi\left(C\sqrt{t}\right) - C\Phi\left(-C\sqrt{t}\right)$$

and

$$\beta_{t,C}(0) = Cq_{tC^2}(0,0) = \frac{1}{\sqrt{t}}\varphi\left(C\sqrt{t}\right) + C\Phi\left(C\sqrt{t}\right).$$

Proof. We will carry out a more detailed proof of the properties on $\beta_{t,C}$. For the case of $\alpha_{t,C}$ the same proof, *mutatis mutandis*, follows as well.

First of all, observe that $\beta_{t,C}(x) = C\beta_{tC^2,1}(Cx)$ and hence it is sufficient to carry out the proof for $C = 1$ then all properties follow for arbitrary $C > 0$.

At the end of the proof of Lemma 4.7 we have shown that the law of $Y_x^-(t)$ coincides with the law of $-Y_{-x}^-(t)$. Hence, the symmetry of $\beta_{t,1}$ follows.

To show analyticity, define $f(s,x) := q_{t-s}(0,0)\rho_{\tau_0^x}(s)$ for $s \in (0,t)$ and $x \in \mathbb{R} \setminus \{0\}$ and the family of domains

$$\mathbb{S}_\varepsilon := \left\{z \in \mathbb{C} :\ \varepsilon < \mathrm{Re}(z) < \frac{1}{\varepsilon},\ \mathrm{Re}(z) > 2|\mathrm{Im}(z)|\right\},$$

$0 < \varepsilon < 1$ and $\mathbb{S} := \cup_{0<\varepsilon<1}\mathbb{S}_\varepsilon$. Then for every $z \in \mathbb{S}$, $g : \mathbb{R}_+ \times \mathbb{S} \to \mathbb{C}$ defined as $g(s,z) := q_{t-s}(0,0)\frac{z}{\sqrt{2\pi s^3}}e^{-\frac{(z-s)^2}{2s}}$ is the holomorphic extension of f to $\mathbb{S}$. Let $\epsilon > 0$, $t > 0$ and let us check that $z \mapsto \int_0^t g(s,z)ds$ is holomorphic on $\mathbb{S}_\varepsilon$. We have $|z| \leq \sqrt{5/4}/\epsilon$, $\mathrm{Re}(z^2) > 3\epsilon^2/4$

and hence

$$
\begin{aligned}
|g(s,z)| &\leq \left(\frac{1}{\sqrt{t-s}}+1\right)\frac{1/\varepsilon}{\sqrt{s^3}}|e^{-\frac{z^2}{2s}}e^{z}e^{-s/2}| \\
&\leq \left(\frac{1}{\sqrt{t-s}}+1\right)\frac{1/\varepsilon}{\sqrt{s^3}}e^{1/\varepsilon}e^{-\frac{3\varepsilon^2}{8s}}
\end{aligned}
$$

for any $s \in (0,t)$, which is integrable on $(0,t)$ for every $\varepsilon > 0$. For a real differentiable function from an open domain in $\mathbb{C}$ to $\mathbb{C}$ we denote the complex conjugate differential operator by $\partial_{\bar{z}}$. Recall, that such a function is holomorphic if and only if its complex conjugate derivative is zero. So, by changing differentiation and integration, we have

$$
\partial_{\bar{z}}\int_0^t g(s,z)ds = \int_0^t \partial_{\bar{z}} g(s,z)ds = 0
$$

for every $z \in \mathbb{S}_\varepsilon$ where the last follows since $g(t,\cdot)$ is holomorphic on $\mathbb{S}$ for every $t > 0$ being thus $\int_0^t f(s,x)ds$ is analytic on $(0,\infty)$. For $x < 0$ use the symmetry of $\beta_{t,1}$ to conclude.

In addition, $\beta_{t,1}$ is Lipschitz in 0, i.e. there is a constant $K > 0$ such that $|\beta_{t,1}(0)-\beta_{t,1}(x)| \leq |x|K$ for any $x \in \mathbb{R}$. Indeed, write

$$
\int_0^t q_{t-s}(0,0)\rho_{\tau_0^x}(s)ds = E[H(\tau_0^x)] + \int_{t/2}^t q_{t-s}(0,0)\rho_{\tau_0^x}(s)(1-h(s))ds
$$

where $H(s) := q_{t-s}(0,0)h(s)$ where h is some function which is bounded by 1, constant 1 near zero, constant 0 on $[t/2,t]$ and $h \in C^\infty([0,t],\mathbb{R})$.

We see that H is Lipschitz continuous with some Lipschitz constant $L > 0$ and, hence,

$$
|\mathrm{E}[H(\tau_0^x)] - \mathrm{E}[H(\tau_0^0)]| \leq L(\mathrm{E}\tau_0^x - \mathrm{E}\tau_0^0) = L|x|
$$

for any $x > 0$. Moreover,

$$
\int_{t/2}^t q_{t-s}(0,0)\rho_{\tau_0^x}(s)(1-h(s))ds \leq |x|\frac{1}{\sqrt{t}}\frac{2}{\pi}\int_{1/2}^1\left(\frac{1}{\sqrt{2\pi t}}\frac{1}{\sqrt{1-s}}+1\right)ds \tag{4.5}
$$

which implies that

$$
|\beta_{t,1}(0)-\beta_{t,1}(x)| \leq |x|K
$$

for some constant $K > 0$. Together with the analyticity outside zero we conclude that $\beta_{t,1}$ is locally Lipschitz continuous. If we have shown that $\beta_{t,1}$ is decreasing on $[0,\infty)$, then it follows that $\beta_{t,1}$ is globally Lipschitz continuous because it is positive valued.

For monotonicity, it is sufficient to show that $\beta_{t,1}$ is decreasing on $(0,\infty)$ and then symmetry and continuity yield the claimed growth properties. Consider $x \in (0,\infty)$ and $v_t^\varepsilon(x) := \mathrm{E}\left[f_\varepsilon(Y_x^-(t))\right]$ where $f_\varepsilon(y) = 1_{\{|y|<\varepsilon\}}$. Here, $\beta_{t,1}(x)$ is defined as the density of $Y_x^-(t)$ at 0. Hence, $\beta_{t,1}(x) = p_t(x,0) = \lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}v_t^\varepsilon(x)$. Thus it is enough to show that $v_t^\varepsilon(x)$ is decreasing on $(0,\infty)$ for every $\varepsilon > 0$. Let $0 < x < y < \infty$. Proposition 4.9 yields $P(\forall t \geq 0 : Y_y^-(t) \geq Y_x^-(t)) = 1$. Define $\tau := \inf\{t > 0 : \ -Y_x^-(t) = Y_y^-(t)\}$. [41, Proposition 2.1.5 a)] yields that τ is a stopping time because it is the first contact time with the closed set $\{0\}$ of the continuous process $Y_x^- + Y_y^-$. Observe, that $|Y_x^-(t)| \leq Y_y^-(t)$ for any $t \in [0,\tau]$. We can write

$$\begin{aligned} v_t^\varepsilon(y) - v_t^\varepsilon(x) &= \mathrm{E}\left[\left(1_{\{|Y_y^-(t)|<\varepsilon\}} - 1_{\{|Y_x^-(t)|<\varepsilon\}}\right) 1_{\{t<\tau\}}\right] \\ &\quad + \mathrm{E}\left[\left(1_{\{|Y_y^-(t)|<\varepsilon\}} - 1_{\{|Y_x^-(t)|<\varepsilon\}}\right) 1_{\{t\geq\tau\}}\right] \\ &= C_1 + C_2 \end{aligned}$$

where $C_1 := \mathrm{E}\left[\left(1_{\{|Y_y^-(t)|<\varepsilon\}} - 1_{\{|Y_x^-(t)|<\varepsilon\}}\right) 1_{\{t<\tau\}}\right]$ and C_2 is the other summand. It can be seen that C_1 is negative since $P(|Y_x^-(t)| \leq \varepsilon, t<\tau) \geq P(|Y_y^-(t)| \leq \varepsilon, t<\tau)$. For the term C_2 we use exactly the same Markov-argument as for the term A_2 in Lemma 4.7 by defining $\tilde{\tau} := \tau \vee t$. Then $\tilde{\tau} \geq \tau$ and $\tilde{\tau}$ is $\mathcal{F}_\tau$-measurable. Thus, the strong Markov property of Y_x^- and Y_y^- and [62, Corollary 2.6.18] yield

$$\begin{aligned} \mathrm{E}\left[1_{\{|Y_y^-(t)|<\varepsilon\}} 1_{\{t\geq\tau\}} | \mathcal{F}_\tau\right] &= \mathrm{E}\left[1_{\{|Y_y^-(\tilde{\tau})|<\varepsilon\}} 1_{\{t\geq\tau\}} | \mathcal{F}_\tau\right] \\ &= 1_{\{t\geq\tau\}} \mathrm{E}\left[1_{\{|Y_y^-(\tilde{\tau})|<\varepsilon\}} | \mathcal{F}_\tau\right] \\ &= 1_{\{t\geq\tau\}} \mathrm{E}\left[1_{\{|Y_y^-(\xi)|<\varepsilon\}} |_{\xi=\tilde{\tau}-\tau}\right] \end{aligned}$$

P-a.s. On the other hand, observe that $Y_y^-(\tau) = -Y_x^-(\tau)$ by the definition of τ. So

$$1_{\{t\geq\tau\}} \mathrm{E}\left[1_{\{|Y_y^-(\xi)|<\varepsilon\}} |_{\xi=\tilde{\tau}-\tau}\right] = 1_{\{t\geq\tau\}} \mathrm{E}\left[1_{\{|-Y_x^-(\xi)|<\varepsilon\}} |_{\xi=\tilde{\tau}-\tau}\right]$$

which implies that $C_2 = 0$. As a result

$$v_t^\varepsilon(y) - v_t^\varepsilon(x) = \mathrm{E}\left[\left(1_{\{|Y_y^-(t)|<\varepsilon\}} - 1_{\{|Y_x^-(t)|<\varepsilon\}}\right) 1_{\{t<\tau\}}\right] \leq 0$$

which implies

$$\beta_t(y) - \beta_t(x) = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}(v_t^\varepsilon(y) - v_t^\varepsilon(x)) \leq 0$$

for every $x, y \in \mathbb{R}$ with $0 < x < y$.

Finally, we show that $\beta_{t,1}$ has exponential tails. Observe that $|q_{t-s}(0,0)| \leq \frac{1}{\sqrt{2\pi(t-t/2)}} + 1$ for $s \in [0, t/2]$ and thus

$$\int_0^{t/2} |q_{t-s}(0,0)| \frac{|x|}{\sqrt{2\pi s^3}} e^{-\frac{(|x|-s)^2}{2s}} \leq K|x|e^{|x|} \int_0^{t/2} s^{-3/2} e^{-\frac{|x|^2}{2s}} ds$$

where K denotes the collection of constants not depending on $x > 0$. Moreover, one can show that

$$\int_0^{t/2} s^{-3/2} e^{-\frac{|x|^2}{2s}} ds \leq K \frac{1}{|x|^2} e^{-\frac{|x|^2}{2t}}$$

for a constant $K > 0$ independent of x. Altogether

$$\int_0^{t/2} |q_{t-s}(0,0)| \frac{|x|}{\sqrt{2\pi s^3}} e^{-\frac{(|x|-s)^2}{2s}} \leq K \frac{e^{|x|}}{|x|} e^{-\frac{|x|^2}{2t}}.$$

Finally, $|\rho_{\tau_0^x}(s)| \leq K|x|e^{-\frac{(|x|-t)^2}{2t}}$ for $s \in [t/2, t]$, $|x| > t$ which yields

$$\int_{t/2}^{t} |q_{t-s}(0,0)||\rho_{\tau_0^x}(s)|ds \leq K|x|e^{-\frac{(|x|-t)^2}{2t}}.$$

□

From now on, let us consider the processes Y_x^- and Y_x^+ given in Equation (4.1) for the multidimensional case, i.e. $x \in \mathbb{R}^d$, $\operatorname{sgn}(x) := \frac{x}{|x|}1_{\{x\neq 0\}}$ and W a d-dimensional standard Brownian motion. We denote $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $W = (W_1, \ldots, W_d)$ and $Y_x^\pm(t) = (Y_{x,1}^\pm(t), \ldots, Y_{x,d}^\pm(t))$. Theorem 4.1 guarantees that the density of any adapted process $X_u(t) := \int_0^t u(s)ds + W(t)$, $u \in \mathcal{A}$, has bounds $\alpha_{d,t} := \alpha_{d,t,1}$ and $\beta_{d,t} := \beta_{d,t,1}$.

We start with a proposition which gives a different view on the functions $\alpha_{d,t,C}$ and $\beta_{d,t,C}$. Namely, we define $Z_x^\pm(t) := |Y_x^\pm(t)|^2$ with $Z_x^\pm(0) = |x|^2$ and denote V_ε the volume of the d-dimensional Euclidean ball of radius ε then we have

$$\alpha_{t,C}(x) = \limsup_{\epsilon\to 0} \frac{P(|Y_x^+(t)| \leq \epsilon)}{V_\epsilon} = \limsup_{\epsilon\to 0} \frac{P(Z_x^+(t) \leq \epsilon^2)}{C_d\ \epsilon^d},$$

and

$$\beta_{t,C}(x) = \limsup_{\epsilon\to 0} \frac{P(|Y_x^-(t)| \leq \epsilon)}{V_\epsilon} = \limsup_{\epsilon\to 0} \frac{P(Z_x^-(t) \leq \epsilon^2)}{C_d\ \epsilon^d},$$

cf. Theorem 4.1, where $C_d := \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$. In view of this equality, we are interested in the behaviour of the transition density of $(Z_x)_{x\in\mathbb{R}}$ near zero which will be exploited in Theorem 4.12 below.

Proposition 4.11. *Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathfrak{A}, P)$ be a filtered probability space. Let W be a d-dimensional Brownian motion and $Y_x^\pm$ be the solution to the SDE*

$$Y_x^\pm(t) := x \pm \int_0^t \operatorname{sgn}(Y_x^\pm(s))ds + W(t), \quad t \geq 0$$

for any $x \in \mathbb{R}^d$. Define $Z_x^\pm(t) := |Y_x^\pm(t)|^2$, $B_x^\pm(t) := \int_0^t \operatorname{sgn}(Y_x^\pm(s))dW(s)$ for any $x \in \mathbb{R}^d$, $t \geq 0$. Then $(Z_x^\pm, B_x^\pm)$ is a solution to the SDE

$$dZ_x^\pm(t) = \left(d \pm 2\sqrt{Z_x^\pm(t)}\right)dt + 2\sqrt{Z_x^\pm(t)}dB_x^\pm(t), \quad Z_x^\pm(0) = |x|^2, \quad t \geq 0 \tag{4.6}$$

for which pathwise uniqueness holds.

Proof. Let $f : \mathbb{R}^d \to \mathbb{R}_+, x \mapsto |x|^2$. Then $Df(x)\cdot y = 2\langle x, y\rangle$ and $H_2 f(x) = 2I_d$ for any $x, y \in \mathbb{R}^d$ where Df denotes the Fréchet differential of f, $H_2 f$ the Hessian matrix of f and I_d denotes the unit matrix in $\mathbb{R}^{d\times d}$. Itô's formula yields

$$\begin{aligned} Z_x^\pm(t) &= |x|^2 + \int_0^t (\pm 2\langle Y_x^\pm(s), \operatorname{sgn}(Y_x^\pm(s))\rangle + d)ds + \int_0^t 2Y_x^\pm(s)dW(s) \\ &= |x|^2 + \int_0^t (\pm 2\sqrt{Z_x^\pm(s)} + d)ds + \int_0^t 2\sqrt{Z_x^\pm(s)}dB_x^\pm(s) \end{aligned}$$

for any $t \geq 0$. Since $B_x^{\pm}$ is a Brownian motion, $(Z_x^{\pm}, B_x^{\pm})$ is a weak solution as required.

It remains to show that the SDE has unique weak solutions. Let Q be a measure, equivalent to P, such that $\widetilde{W}^{\pm}(t) := B_x^{\pm}(t) - t$ is a standard Q-Brownian motion. Then the SDE can be rewritten as

$$dZ_x^{\pm}(t) = (d)\,dt \pm 2\sqrt{Z_x^{\pm}(t)}d\widetilde{W}^{\pm}(t), \quad Z_x^{\pm}(0) = |x|^2, \quad t \geq 0. \tag{4.7}$$

[100, Theorem IX.3.5 ii)] yields that pathwise uniqueness holds for SDE (4.7) under Q. □

The following result gives explicit bounds for the functions $\alpha_{d,t}$ and $\beta_{d,t}$.

Theorem 4.12. *We have*

$$\frac{2^d}{C_d d^{d/2}} \prod_{i=1}^{d} \alpha_{1,t}(x_i) \leq \alpha_{d,t}(x) \leq \beta_{d,t}(x) \leq \frac{2^d}{C_d} \prod_{i=1}^{d} \beta_{1,t}(x_i), \quad x \in \mathbb{R}^d$$

where $C_d := \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$.

Proof. Since the proof is fairly similar for $\alpha_{d,t}$, we will just show the last inequality.

Define the processes $Z_{x,i}^{-}(t) := |Y_{x,i}^{-}(t)|^2$, $i = 1, \ldots, d$. Itô's formula yields

$$\begin{aligned} Z_{x,i}^{-}(t) &= |x_i|^2 + \int_0^t \left(1 - 2\sqrt{Z_{x,i}^{-}(s)}\frac{|Y_{x,i}^{-}(s)|}{|Y_x^{-}(s)|}\right) ds + 2\int_0^t Y_{x,i}^{-}(s)dW_i(s) \\ &= |x_i|^2 + \int_0^t \left(1 - 2\sqrt{Z_{x,i}^{-}(s)}\frac{|Y_{x,i}^{-}(s)|}{|Y_x^{-}(s)|}\right) ds + 2\int_0^t \sqrt{Z_{x,i}^{-}(s)}dB_i(s) \\ &\geq |x_i|^2 + \int_0^t \left(1 - 2\sqrt{Z_{x,i}^{-}(s)}\right) ds + 2\int_0^t \sqrt{Z_{x,i}^{-}(s)}dB_i(s) \end{aligned}$$

where $B_i(t) := \int_0^t \operatorname{sgn}(Y_{x,i}^{-}(s))dW_i(s)$ defines a new standard Brownian motion w.r.t. P. [100, Theorem IV.3.6] and Itô isometry ensure that $(B_1, \ldots, B_d)$ is a d-dimensional standard Brownian motion. Let V_i be the solution of the SDE

$$V_i(t) = |x_i|^2 + \int_0^t \left(1 - 2\sqrt{V_i(s)}\right) ds + 2\int_0^t \sqrt{V_i(s)}dB_i(s) \tag{4.8}$$

for any $i = 1, \ldots, d$ and Q be the measure, equivalent to P, such that $\widetilde{B}(t) := B(t) - (t, \ldots, t)$ is a Q-Brownian motion where $B = (B_1, \ldots, B_d)$. Then, we have

$$\begin{aligned} Z_{x,i}^{-}(t) &= |x_i|^2 + \int_0^t \left(1 + 2\sqrt{Z_{x,i}^{-}(s)}\left(1 - \frac{|Y_{x,i}^{-}(s)|}{|Y_x^{-}(s)|}\right)\right) ds + 2\int_0^t \sqrt{Z_{x,i}^{-}(s)}d\widetilde{B}_i(s), \\ V_i(t) &= |x_i|^2 + \int_0^t 1 ds + 2\int_0^t \sqrt{V_i(s)}d\widetilde{B}_i(s). \end{aligned}$$

Similar arguments as in the proof of [100, Theorem IX.3.7] show that $Z_{x,i}^{-}(t) \geq V_i(t)$ for any $t \geq 0$, Q-a.s.

Observe that pathwise uniqueness holds for Equation (4.6) by Proposition 4.11 and hence [100, Theorem IX.1.7 ii)] states that V_i is a strong solution to Equation (4.8). Consequently, V_i is $\sigma(B_i)$-measurable and hence $V_1, \dots, V_d$ are independent processes.

Now given $a = (a_1, \dots, a_d) \in \mathbb{R}^d$ one has $|a| \geq \max\{|a_i|, i = 1, \dots, d\}$. This implies

$$\begin{aligned} P(|Y_x^-(t)| \leq \varepsilon) &\leq P\left(\bigcap_{i=1}^d \{|Y_{x,i}^-(t)| \leq \varepsilon\}\right) \\ &= P\left(\bigcap_{i=1}^d \{Z_{x,i}^-(t) \leq \varepsilon^2\}\right) \\ &\leq \prod_{i=1}^d P\left(V_i(t) \leq \varepsilon^2\right) \end{aligned}$$

where in the last step we use the inequalities $Z_{x,i}^-(t) \geq V_i(t)$ for every $t \geq 0$, P-a.s. and the fact that $V_1, \dots, V_d$ are independent processes.

By Proposition 4.11 the law of $V_i(t)$ under P is the same as the law of $|A_i(t)|^2$ under P where

$$A_i(t) = x_i - \int_0^t \operatorname{sgn}(A_i(s))ds + W_i(t), \quad t \geq 0$$

and the law of $A_i(t)$ is given in Lemma 4.7. Hence, we have

$$\begin{aligned} \beta_{d,t}(x) &\xleftarrow{\varepsilon \to 0} \frac{P(|Y_x^-(t)| \leq \varepsilon)}{C_d \varepsilon^d} \\ &\leq \frac{1}{C_d} \prod_{i=1}^d \frac{P(|A_i(t)| \leq \varepsilon)}{\varepsilon} \\ &\xrightarrow{\varepsilon \to 0} \frac{2^d}{C_d} \prod_{i=1}^d \beta_{1,t}(x_i) \end{aligned}$$

for any $t > 0$. □

In order to prepare our main result of this section we will start with a series of lemmas which aims at showing the continuity condition of [60, Theorem IX.2.11]. The needed continuity condition is summarised in Lemma 4.16.

Lemma 4.13. *Let $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathfrak{A}, P)$ be a filtered probability space. Let $\varphi : \mathbb{R}_+ \to [0, 1]$ such that φ is infinitely differentiable, φ is constant 1 on $[0, 1]$ and constant 0 on $[2, \infty)$. Define*

$$A_k : \mathbb{R}_+ \times \mathbb{D}(\mathbb{R}^d) \to \mathbb{R}, (t, f) \mapsto \int_0^t \varphi(k|f(s)|)ds$$

for any $k \in \mathbb{N}$. Let $b : \Omega \times \mathbb{R}_+ \to \mathbb{R}^d$ be an adapted process which is bounded by 1, $x \in \mathbb{R}^d$ and define

$$X(t) := x + \int_0^t b(s)ds + W(t), \quad t \geq 0.$$

Then $\mathrm{E}(A_k(t, X)) \leq \sqrt{tc_k(t)} \exp(t/2)$ where $c_k(t) := \mathrm{E}(A_k(t, x + W)) \to 0$ for $k \to \infty$.

Proof. Define $Z(t) := \mathcal{E}(-\int_0^t b(s)dW(s)) = 1 - \int_0^t Z(s)b(s)dW(s)$ and $dQ|_{\mathcal{F}_t} := Z_t dP|_{\mathcal{F}_t}$. Then Girsanov's theorem [60, Theorem III.3.24] yields that X is a Q-Brownian motion starting in x. Define $Y(t) := 1/Z(t)$. Then

$$Y(t) = 1 + \int_0^t Y(s)b(s)dX(s), \quad t \geq 0$$

and hence by Grönwall's lemma, see e.g. [100, Appendix §1], $\mathrm{E}_Q(Y(t)^2) \leq \exp(t)$. We have

$$\begin{aligned}\mathrm{E}(A_k(t,X)) &= \mathrm{E}_Q\left(A_k(t,X)Y(t)\right) \\ &\leq \sqrt{\mathrm{E}_Q(A_k(t,X)^2)}\sqrt{\mathrm{E}_Q(Y^2(t))} \\ &\leq \sqrt{t\mathrm{E}_Q(A_k(t,X))}\exp(t/2) \\ &= \sqrt{tc_k(t)}\exp(t/2)\end{aligned}$$

for any $t \geq 0$ where we used the Cauchy-Schwartz inequality twice and the fact that $\varphi^2 \leq \varphi$. We have

$$\begin{aligned}c_k(t) &= \mathrm{E}(A_k(t, x + W)) \\ &\to \mathrm{E}(\lambda(\{s \in [0,t] : x + W(s) = 0\}) \\ &= 0\end{aligned}$$

for $k \to \infty$ by Lebesgue's dominated convergence theorem where λ denotes the Lebesgue measure on $\mathbb{R}$. □

Lemma 4.14. *Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathfrak{A}, P)$ be a filtered probability space and W be a d-dimensional standard Brownian motion. Let $(A_k)_{k\in\mathbb{N}}$ and $(c_k)_{k\in\mathbb{N}}$ be as in Lemma 4.13. Let $(M_n)_{n\in\mathbb{N}}$ be a sequence of processes that converges in probability to W. For any $n \in \mathbb{N}$ let $b_n : \Omega \times \mathbb{R}_+ \to \mathbb{R}^d$ be an adapted process which is bounded by 1, $x \in \mathbb{R}^d$ and define*

$$X_n(t) := x + \int_0^t b_n(s)ds + M_n(t), \quad t \geq 0.$$

Also, assume that $(X_n)_{n\in\mathbb{N}}$ converges in distribution to some process X_∞.

Then X_∞ has P-a.s. continuous sample paths and

$$\mathrm{E}A_k(t, X_\infty) \leq \sqrt{tc_k(t)}\exp(t/2), \quad t \geq 0, \quad k \in \mathbb{N}.$$

Moreover, $\lambda\{s \in \mathbb{R}_+ : X_\infty(s) = 0\} = 0$ P-a.s. where λ denotes the Lebesgue measure on $\mathbb{R}$.

Proof. Define $Y_n(t) := x + \int_0^t b_n(s)ds + W(t)$, $t \geq 0$. Then

$$X_n - Y_n = M_n - W \to 0$$

in probability for $n \to \infty$. Hence, $Y_n \to X_\infty$ in distribution. Since Y_n has continuous sample paths for any $n \in \mathbb{N}$, X_∞ has P-a.s. continuous sample paths. Let $t, \epsilon > 0$. Since A_k is continuous we have $\mathrm{E}A_k(t, X_\infty) \leq \epsilon + \mathrm{E}A_k(t, Y_n)$ for some $n \in \mathbb{N}$. Hence, by Lemma 4.13

we have

$$\begin{aligned}\mathrm{E}A_k(t, X_\infty) &\leq \epsilon + \mathrm{E}A_k(t, Y_n)\\ &\leq \epsilon + \sqrt{tc_k(t)}\exp(t/2).\end{aligned}$$

Thus, we have

$$\begin{aligned}\mathrm{E}(\lambda\{s \in [0,t] : X_\infty(s) = 0\}) &\leftarrow \mathrm{E}A_k(t, X_\infty)\\ &\to 0\end{aligned}$$

for $k \to \infty$ and $t \geq 0$. Thus $\mathrm{E}(\lambda\{s \in \mathbb{R}_+ : X_\infty(s) = 0\}) \leq \sum_{n=1}^{\infty} \mathrm{E}(\lambda\{s \in [0,n] : X_\infty(s) = 0\}) = 0$. The claim follows. □

Remark 4.15. *Let $x \in \mathbb{R}^d \backslash \{0\}$, $\epsilon \in (0, |x|)$ and $y \in \mathbb{R}^d$ such that $|x - y| \leq \epsilon$. Then*

$$|\mathrm{sgn}(x) - \mathrm{sgn}(y)| \leq \sqrt{2}\left(\frac{\epsilon}{|x|}\right).$$

Lemma 4.16. *Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathfrak{A}, P)$ be a filtered proability space and W be a d-dimensional standard Brownian motion. Let $(b_n)_{n\in\mathbb{N}}$ be adapted processes which are bounded by 1. Let $x \in \mathbb{R}^d$, $(M_n)_{n\in\mathbb{N}}$ be a sequence of adapted processes which converges in probability to W and define*

$$X_n(t) := x + \int_0^t b_n(s)ds + M_n(t).$$

Assume that $(X_n)_{n\in\mathbb{N}}$ converges in distribution to some process X_∞ and define

$$B : \mathbb{R}_+ \times \mathbb{D}(\mathbb{R}^d) \to \mathbb{R}^d, (t, f) \mapsto -\int_0^t \mathrm{sgn}(f(s))ds, \quad t \geq 0$$

Then $f \mapsto B(t, f)$ is P^{X_∞}-a.s. continuous for any $t \geq 0$.

Proof. Let A_k be as in Lemma 4.13 for any $k \in \mathbb{N}$. Lemma 4.14 yields

$$\lambda\{s \in \mathbb{R}_+ : X_\infty(s) = 0\} = 0$$

P-a.s. and hence $A_k(X_\infty, t) \to 0$ for $k \to \infty$ P-a.s.

Let $t \geq 0$ and $f, g_k \in \mathbb{D}(\mathbb{R}^d)$ such that $\sup\{|f(s) - g_k(s)| : s \leq t\} \leq 1/k^2$ for any $k \in \mathbb{N}$.

Then, we have

$$
\begin{aligned}
|B(t,f)-B(t,g_k)| \leq &\int_0^t |\mathrm{sgn}(f(s))-\mathrm{sgn}(g_k(s))|ds\\
=&\int_0^t |\mathrm{sgn}(f(s))-\mathrm{sgn}(g_k(s))|1_{\{|f(s)|\leq 1/k\}}ds\\
&+\int_0^t |\mathrm{sgn}(f(s))-\mathrm{sgn}(g_k(s))|1_{\{|f(s)|>1/k\}}ds\\
\leq &2\int_0^t 1_{\{|f(s)|\leq 1/k\}}ds + t\sqrt{2}/k\\
\leq &2\int_0^t \varphi(k|f(s)|)ds + t\sqrt{2}/k\\
=&2A_k(t,f)+t\sqrt{2}/k\\
\to &0
\end{aligned}
$$

P^{X_∞}-a.s. for $k \to \infty$ where we used the integral inequality, then we split the support of f, Remark 4.15 with $\epsilon = 1/k^2$ and the inequality $1_{[0,1]}(x) \leq \varphi(x)$ for any $x \geq 0$. □

In the next lemma the martingales M_n converge to the Brownian motion W but they, and hence the drift in X_n, are not adapted to the same Brownian motion. We show that they converge in our specific set-up.

Lemma 4.17. *Let $(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathfrak{A}, P)$ be a filtered probability space. Let W be a d-dimensional Brownian motion, $x \in \mathbb{R}^d$ and $M_n(t) := W(\theta_n(t))$ where $\theta_n(t) := \inf\{k/n : t < k/n\}$ for any $n \in \mathbb{N}$. Assume that $X_n(t) = x - \int_0^t \mathrm{sgn}(X_n(s))ds + M_n(t)$ for any $n \in \mathbb{N}$. Then $(X_n)_{n\in\mathbb{N}}$ converges in distribution to the solution X of the SDE*

$$X(t) = x - \int_0^t \mathrm{sgn}(X(s))ds + W(t), \quad t \geq 0. \tag{4.9}$$

Proof. By an independent enlargement of $\mathcal{F}_0$-argument, we may assume that there is a sequence $(H_n)_{n\in\mathbb{N}}$ of random variables which are indepent of W, $\mathcal{F}_0$-measurable and that H_n is centered normal on $\mathbb{R}^d$ with variance I_d/n where I_d denotes the identity matrix in $\mathbb{R}^{d\times d}$.

Define $\tilde{\theta}_n(t) := \theta_n(t) - 1/n = \max\{k/n : k \geq 0, k/n \leq t\}$ for any $n \in \mathbb{N}$. Then $0 \leq \tilde{\theta}_n(t) \leq t$. Define the $\mathcal{F}$-adapted process $\widetilde{M}_n(t) := H_n + W(\widetilde{\theta}_n(t))$ and $\widetilde{X}_n(t) = x - \int_0^t \mathrm{sgn}(\widetilde{X}_n(s))ds + \widetilde{M}_n(t)$. Then M_n has the same law as $\widetilde{M}_n$ and, consequently, (X_n, M_n) has the same law as $(\widetilde{X}_n, \widetilde{M}_n)$ for any $n \in \mathbb{N}$. Moreover, $\widetilde{M}_n \to W$ in probability.

Define

$$\begin{aligned}
B(t,f) &:= -\int_0^t \operatorname{sgn}(f(s))ds,\\
C(t,f) &:= tI_d,\\
\nu(A\times I) &:= 0,\\
B_n(t) &:= B(t,\widetilde{X}_n),\\
C_n(t) &:= 0I_d = 0,\\
\nu_n(A\times I) &:= \mu_n(A)\sum_{k=1}^{\infty}\delta_{k/n}(I)
\end{aligned}$$

for any $t\in\mathbb{R}_+$, $f\in\mathbb{D}(\mathbb{R}^d)$, $n\in\mathbb{N}$, $A\in\mathcal{B}(\mathbb{R}^d)$, $I\in\mathcal{B}(\mathbb{R}_+)$ where μ_n is the centered normal law with covariance matrix I_d/n. Then (B_n,C_n,ν_n) is the semimartingale characteristics of X_n in the sense of [60, Definition II.2.6] relative to the truncation function $h(x):=\operatorname{sgn}(x)(|x|\wedge 1)$, $x\in\mathbb{R}^d$. Observe that $(B_n,C_n,\nu_n)_{n\in\mathbb{N}}$ and (B,C,ν) fulfil the conditions $[\mathrm{Sup}{-}\beta_7]$, $[\mathrm{Sup}{-}\gamma_7]$ and $[\mathrm{Sup}{-}\delta_{7,1}]$ in the sense of [60, page 535]. Thus [60, Theorem IX.3.9] states that $(\widetilde{X}_n)_{n\in\mathbb{N}}$ is tight. Let $(\widetilde{X}_{n_j})_{j\in\mathbb{N}}$ be a subsequence of $(\widetilde{X}_n)_{n\in\mathbb{N}}$ which converges in law and denote the limiting law by P_∞. Lemma 4.16 yields that B is P_∞-a.s. continuous. Let Y be the canonical process on the canonical space $(\mathbb{D}(\mathbb{R}^d),(\mathcal{G}_t)_{t\geq 0},\mathcal{B}(\mathbb{D}(\mathbb{R}^d)))$. Then [60, Theorem IX.2.11] yields that Y is, under P_∞, a semimartingale with characteristics (B,C,ν). The continuous martingale part, denote it $\widetilde{W}$, of Y is a standard Brownian motion because its semimartingale characteristics is $(0,C,0)$. Moreover,

$$Y(t)=x+B(t,Y)+\widetilde{W}(t)=x-\int_0^t \operatorname{sgn}(Y(s))ds+\widetilde{W}(t),\quad t\geq 0.$$

Thus $(Y,\widetilde{W})$ is a weak solution to the SDE (4.9). [100, Corollary IX.1.12)] yields that the law P_∞ of Y coincides with the law of the solution X of the SDE (4.9). Consequently, any convergent subsequence of $(\widetilde{X}_n)_{n\in\mathbb{N}}$ converges in law to X. Since $(\widetilde{X}_n)_{n\in\mathbb{N}}$ is additionally tight, it, and hence $(X_n)_{n\in\mathbb{N}}$, converges to X. □

We are now in readiness to prove the core result of this section which is to solve a control problem. For dimension one, this problem has been studied by V. E. Beneš in [17] in the Markovian setting whose optimal control is indeed the *signum* function in dimension one. Such solutions are known as *bang-bang* solutions. Nevertheless, here we stress the fact that our case deals with multivalued controls, thus not *bang-bang*, in addition to the non-Markovian setting as in the above mentioned result. For this reason we include a short proof based on the previous limit results.

Theorem 4.18. *Let $\mathcal{A}_+$ and $\mathcal{A}$ be as in the beginning of Section 4.2. Let $T,\epsilon>0$, $x\in\mathbb{R}^d$ and define $u_x^*(t):=\operatorname{sgn}(Y_x^+(t))$ and $v_x^*(t):=-\operatorname{sgn}(Y_x^-(t))$. Then*

$$\inf_{u\in\mathcal{A}} P(|X_u(T)|\leq\epsilon)=P(|X_{u_x^*}(T)|\leq\epsilon)\tag{4.10}$$

where $X_u(t) := x + \int_0^t u(s)ds + W(t)$ for $u \in \mathcal{A}$. In other words, an optimal control for the control problem above is given by u_x^*. Similarly,

$$\sup_{v\in\mathcal{A}} P(|X_v(T)| \leq \epsilon) = P(|X_{v_x^*}(T)| \leq \epsilon). \tag{4.11}$$

Remark 4.19. *The control problem given in (4.10) can be interpreted as follows: one wishes to find the stochastic process among those in $\mathcal{A}$ that minimises the probability that the underlying process X is near zero. In other words, we want the process $X_u(T)$ to escape from 0 as much as possible. Intuitively, the process Y_x^+ is doing that. Whenever $Y_x^+(t)$ is near zero on the positive line, the drift $\mathrm{sgn}(Y_x^+(t))$ is positive and pushes $Y_x^+(t)$ even further away up and if $Y_x^+(t)$ is near zero from below the drift is negative and sends $Y_x^+(t)$ further down. For the control problem in (4.11) the idea is similar, but there one wishes to maximise the probability of being close to zero, which $-\mathrm{sgn}(Y_x^-(t))$ clearly does.*

For a general reference on control problems we relate to Øksendal and Sulem [96].

Proof of Theorem 4.18. For the sake of brevity we will only show the proof of the control for (4.11).

For any $n \in \mathbb{N}$ define $\theta_n(t) := \inf\{Tk/n : k \in \mathbb{N}, t < Tk/n\}$, $M_n(t) := W(\theta_n(t))$ and

$$\mathcal{A}_n := \{v \in \mathcal{A}_+ : v(t) \text{ is } \mathcal{F}_{\theta_n(t)}\text{-measurable for any } t \in [0,T]\}$$

Then M_n is adapted to the filtration $(\mathcal{G}_{n,t})_{t\geq 0} := (\mathcal{F}_{\theta_n(t)})_{t\geq 0}$.

Let $X_n(t) = x - \int_0^t \mathrm{sgn}(X_n(s))ds + M_n(t)$, $t \geq 0$. A simple backward induction yields that

$$P(|X_n(T-)| \leq \epsilon) = \sup_{v\in\mathcal{A}_n} P\left(\left|x + \int_0^T v(s)ds + M_n(T-)\right| \leq \epsilon\right).$$

Lemma 4.17 yields that $(X_n)_{n\in\mathbb{N}}$ converges in law to Y_x^-. Since $Y_x^-(T)$ has no atoms, we have $P(|X_n(T)| \leq \epsilon) \to P(|Y_x^-(T)| \leq \epsilon)$ for $n \to \infty$. Thus, we have

$$\begin{aligned}
P(|Y_x^-(T)| \leq \epsilon) &\leq \sup_{v\in\mathcal{A}} P(|X_v(T)| \leq \epsilon) \\
&\leq \sup_{v\in\mathcal{A}_n} P(|X_v(T)| \leq \epsilon) \\
&\leq \sup_{v\in\mathcal{A}_n} P\left(\left|x + \int_0^T v(s)ds + M_n(T-)\right| \leq \epsilon\right) \\
&= P(|X_n(T-)| \leq \epsilon) \\
&\to P(|Y_x^-(T)| \leq \epsilon)
\end{aligned}$$

for $n \to \infty$. Thus v_x^* is an optimal control.

□

Finally, we give the proof of our main result Theorem 4.1.

Proof of Theorem 4.1. Define $\widetilde{X}(t) := CX(t/C^2)$, $\widetilde{u}(t) := u(t/C^2)$ and the Brownian motion

$\widetilde{W}(t) := CW(t/C^2)$. Then

$$\begin{aligned}\widetilde{X}(t) &= \int_0^{t/C^2} C^2 u(s)ds + \widetilde{W}(t)\\ &= \int_0^t \widetilde{u}(s)ds + \widetilde{W}(t)\end{aligned}$$

for any $t \geq 0$. Theorem 4.18 states that

$$P(|\widetilde{X}(T) + x| \leq \epsilon) \leq P(|Y_x^-(T)| \leq \epsilon)$$

for any $\epsilon, T > 0$, $x \in \mathbb{R}^d$ and $u \in \mathcal{A}$. By definition

$$\lim_{\epsilon \to 0} \frac{P(|Y_x^-(T)| \leq \epsilon)}{V_\epsilon} = \beta_{d,T,1}(x).$$

Thus we have

$$\rho_{C,T}(x) := \limsup_{\epsilon \to 0} \frac{P(|\widetilde{X}(T) - x| \leq \epsilon)}{V_\epsilon} \leq \beta_{d,T,1}(-x).$$

Observe that for any orthonormal transformation $U : \mathbb{R}^d \to \mathbb{R}^d$ we have

$$UY_x^-(t) = Ux - \int_0^t \operatorname{sgn}(UY_x^-(s))ds + UW(t)$$

where here UW is a standard Brownian motion and hence (UY_x^-, UW) is a weak solution of (4.1) for $\pm = -$. Consequently, $UY_x^-(t)$ has the same law as Y_{Ux}^- which implies $\beta_{d,T,1}(Ux) = \beta_{d,T,1}(x)$. Hence, we have

$$\rho_{C,T}(x) \leq \beta_{d,T,1}(x).$$

Lebesgue differentiation theorem [53, Corollary 2.1.16] yields that $\rho_{C,T}$ is a version of the Lebesgue density of $\widetilde{X}(T)$. Consequently, the density ρ_T of $X(T)$ given by

$$\rho_T(x) := \limsup_{\epsilon \to 0} \frac{P(|X(T) - x| \leq \epsilon)}{V_\epsilon}$$

satisfies

$$\rho_T(x) \leq \beta_{d,T,C}(x).$$

Analogue arguments show that

$$\alpha_{d,T,C}(x) \leq \rho_T(x).$$

□

Chapter 5

Construction of Malliavin differentiable strong solutions of SDEs under an integrability condition on the drift without the Yamada-Watanabe principle

David R. Baños, Sindre Duedahl, Thilo Meyer-Brandis and Frank Proske

Abstract: In this paper we aim at employing a compactness criterion of Da Prato, Malliavin, Nualart [30] for square integrable Brownian functionals to construct unique strong solutions of SDE's under an integrability condition on the drift coefficients. The obtained solutions turn out to be Malliavin differentiable and are used to derive a Bismut-Elworthy-Li formula for solutions of the Kolmogorov equation.

5.1 Introduction

The object of study of this paper is the stochastic differential equation (SDE)

$$X_t = x + \int_0^t b(s, X_s^x)ds + B_t, \quad 0 \leq t \leq T, \quad x \in \mathbb{R}^d, \tag{5.1}$$

where $B.$ is a d-dimensional Brownian motion on some complete probability space $(\Omega, \mathcal{F}, \mu)$ with respect to a μ-completed Brownian filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ and where $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a Borel-measurable function.

In this article we are interested in the analysis of strong solutions $X.$ of the SDE (5.1), that is an $\{\mathcal{F}_t\}_{0 \leq t \leq T}$-adapted solution processes on $(\Omega, \mathcal{F}, \mu)$ when the drift coefficient is irregular, e.g. non-Lipschitzian or discontinuous.

A widely used construction method for strong solutions in this case in the literature is based on the so-called Yamada-Watanabe principle. Using this principle, a once constructed weak solution, that is a solution which is not necessarily a functional of the driving noise, combined with pathwise uniqueness gives a unique strong solution. So

$$\boxed{\text{Weak solution}} \quad + \quad \boxed{\text{Pathwise uniqueness}} \quad \Rightarrow \quad \boxed{\text{Unique strong solution}}. \tag{5.2}$$

Here, pathwise uniqueness means the following: If $X_\cdot^{(1)}$ and $X_\cdot^{(2)}$ are $\{\mathcal{F}_t^{(1)}\}_{0\leq t\leq T}$- and respectively $\{\mathcal{F}_t^{(2)}\}_{0\leq t\leq T}$-adapted weak solutions on a probability space, then these solutions must coincide a.s. See [108]. In the milestone paper from 1974 [110], A.K. Zvonkin used the Yamada-Watanabe principle in the one-dimensional case in connection with PDE techniques to construct a unique strong solution to (5.1), when b is merely bounded and measurable. Subsequently, the latter result was generalised by A.Y. Veretennikov [105] to the multidimensional case.

Important other and more recent results in this direction are e.g. [68], [54] and [67]. See also the striking work [28] in the Hilbert space setting, where the authors use solutions of infinite-dimensional Kolmogorov equations to obtain unique strong solutions of stochastic evolution equations with bounded and measurable drift for a.e. initial values.

In this article we want to employ a construction principle for strong solutions developed in [83]. This method which relies on a compactness criterion from Malliavin Calculus for square integrable functionals of the Brownian motion [30] is in diametrical opposition to the Yamada-Watanabe principle (5.2) in the sense that

$$\boxed{\text{Strong existence}} \quad + \quad \boxed{\text{Uniqueness in law}} \quad \Rightarrow \quad \boxed{\text{Strong uniqueness}},$$

that is the existence of a strong solution to (5.1) and uniqueness in law of solutions imply the existence of a unique strong solution. A crucial consequence of this approach is the additional insight that the constructed solutions are regular in the sense of Malliavin differentiability.

We mention that this method has been recently applied in a series of other papers. See e.g. [81], where the authors obtain Malliavin differentiable solutions when the drift coefficient in $\mathbb{R}^d$ is bounded and measurable. Other applications pertain to the stochastic transport equation with singular coefficients [87], [88] or stochastic evolution equations in Hilbert spaces with bounded Hölder-continuous drift [48]. See also [56] in the case of truncated α-stable processes as driving noise and [15] in the case of fractional Brownian motion for Hurst parameter $H < 1/2$, which is a non-Markovian driving noise.

Using the above mentioned new approach, one of the objectives of this paper is to construct Malliavin differentiable unique strong solutions to (5.1) under the integrability condition

$$b \in L^q([0,T], L^p(\mathbb{R}^d, \mathbb{R}^d)) \tag{5.3}$$

for $p \geq 2$, $q > 2$ such that

$$\frac{d}{p} + \frac{2}{q} < 1.$$

The idea for the proof rests on a mixture of techniques in [81] and [44]. More precisely, we approximate in the first step the drift coefficient b by smooth functions b_n with compact support and apply the Itô-Tanaka-Zvonkin "trick" by transforming the solutions $X_t^{n,x}$ of (5.1) associated with the coefficients b_n to processes

$$Y_t^{n,x} := X_t^{n,x} + U_n(t, X_t^{n,x}),$$

where the processes $Y_t^{n,x}$ satisfy an equation with more regular coefficients than (5.1) given by

$$dY_t^{n,x} = \lambda U_n(t, X_t^{n,x})dt + (\mathcal{I}_d + \nabla U_n(t, X_t^{n,x}))\, dB_t$$

for solutions U_n to the backward PDE's

$$\frac{\partial U_n}{\partial t} + \frac{1}{2}\Delta U_n + b_n \nabla U_n = \lambda U_n - b_n, \quad U_n(T, x) = 0. \tag{5.4}$$

In the second step we use the compactness criterion for $L^2(\Omega)$ in [30] applied to the sequence $Y_t^{n,x}$, $n \geq 1$ in connection with Schauder-type of estimates of solutions of (5.4) and techniques from white noise analysis to show that

$$Y_t^{n,x} \xrightarrow{n\to\infty} Y_t^x$$

in $L^2(\Omega)$ for all t and that

$$X_t^x = \varphi(t, Y_t^x),$$

where $\varphi(t, \cdot)$ is the inverse of the function $x \mapsto x + U(t, x)$ for all t and U a solution of (5.4), is a Malliavin differentiable unique strong solution of (5.1).

Our paper is organised as follows: In Section 5.2 we present our main results on the construction of strong solutions (Theorem 5.1 and Theorem 5.15). As an application of the results obtained in Section 5.2 we establish in Section 5.3 a Bismut-Elworthy-Li formula for the representation of first order derivatives of solutions of Kolmogorov equations.

5.2 Main results

In this section, we want to further develop the ideas introduced in [44] and [83] to derive Malliavin differentiable strong solutions of stochastic differential equations with irregular coefficients. More precisely, we aim at analysing the SDE's of the form

$$dX_t = b(t, X_t)dt + dB_t,\ 0 \leq t \leq 1,\ \ X_0 = x \in \mathbb{R}^d\,, \tag{5.5}$$

where the drift coefficient $b : [0, T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a Borel measurable function satisfying some integrability condition and B_t is a d-dimensional Brownian motion with respect to the stochastic basis

$$(\Omega, \mathcal{F}, \mu)\,, \{\mathcal{F}_t\}_{0\leq t\leq T} \tag{5.6}$$

for the $\mu-$augmented filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ generated by B_t. At the end of this section we shall also apply our technique to equations with more general diffusion coefficients (Theorem 5.15).

Consider the space

$$L_p^q := L^q\left([0, T], L^p(\mathbb{R}^d, \mathbb{R}^d)\right)$$

for $p, q \in \mathbb{R}$ satisfying the following condition

$$p > 2,\ q > 2 \ \text{ and } \ \frac{d}{p} + \frac{2}{q} < 1 \tag{5.7}$$

and denote by $|\cdot|$ the Euclidean norm in $\mathbb{R}^d$. The Banach space L_p^q is endowed with the norm

$$\|f\|_{L_p^q} = \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(t,x)|^p dx \right)^{q/p} dt \right)^{1/q} < \infty \tag{5.8}$$

for $f \in L_p^q$.

The main goal of the paper is to show that SDE's of the type (5.5) with drift coefficient b satisfying the integrability condition given in (5.8) admit strong solutions that are unique and in addition, Malliavin differentiable.

So, our main result is the following theorem:

Theorem 5.1. *Suppose that the drift coefficient $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ in* (5.5) *belongs to L_p^q. Then there exists a unique global strong solution X to equation* (5.5) *such that X_t is Malliavin differentiable for all $0 \le t \le T$.*

An important step of the proof of Theorem 5.1 is directly based on the study of the regularity of solutions to the following associated PDE to equation SDE (5.5).

$$\partial_t U(t,x) + b(t,x) \cdot \nabla U(t,x) + \frac{1}{2}\Delta U(t,x) - \lambda U(t,x) + b = 0,\ \ t \in [0,T],\ \ U(T,x) = 0, \tag{5.9}$$

where $U : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\lambda > 0$ and $b \in L_p^q$.

The following result is due to [44] and stablishes the well-posedness of the above PDE problem in a certain space.

First, recall the definition of the following functional spaces

$$\mathbb{H}_{\alpha,p}^q = L^q([0,T], W^{\alpha,p}(\mathbb{R}^d)),\ \ \mathbb{H}_p^{\beta,q} = W^{\beta,q}([0,T], L^p(\mathbb{R}^d))$$

and

$$H_{\alpha,p}^q = \mathbb{H}_{\alpha,p}^q \cap \mathbb{H}_p^{1,q}.$$

The norm in $H_{\alpha,p}^q$ can be taken to be

$$\|u\|_{H_{\alpha,p}^q} \equiv \|u\|_{\mathbb{H}_{\alpha,p}^q} + \|\partial_t u\|_{L_p^q}.$$

Theorem 5.2. *Let p, q be such that $p \ge 2$, $q > 2$ and $\frac{d}{p} + \frac{2}{q} < 1$ and $\lambda > 0$. Consider two vector fields $b, \Phi \in L_p^q$. Then there exists a unique solution of the backward parabolic system*

$$\partial_t u + \frac{1}{2}\Delta u + b \cdot \nabla u - \lambda u + \Phi = 0,\ \ t \in [0,T],\ \ u(T,x) = 0 \tag{5.10}$$

belonging to the space

$$H^q_{2,p} := L^q([0,T], W^{2,p}(\mathbb{R}^d)) \cap W^{1,q}([0,T], L^p(\mathbb{R}^d)),$$

i.e. there exists a constant $C > 0$ depending only on d, p, q, T, λ and $\|b\|_{L^q_p}$ such that

$$\|u\|_{H^q_{2,p}} \leq C\|\Phi\|_{L^q_p}. \tag{5.11}$$

The following result is a part of [68, Lemma 10.2] that gives us some properties on the regularity of $u \in H^q_{2,p}$ that we will need for the proof of Theorem 5.1.

Lemma 5.3. *Let $p, q \in (1,\infty)$ such that $\frac{d}{p} + \frac{2}{q} < 1$ and $u \in H^q_{2,p}$, then ∇u is Hölder continuous in $(t,x) \in [0,T] \times \mathbb{R}^d$, namely for any $\varepsilon \in (0,1)$ satisfying*

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1$$

there exists a constant $C > 0$ depending only on p, q and ε such that for all $s, t \in [0,T]$ and $x, y \in \mathbb{R}^d$, $x \neq y$

$$\|\nabla u(t,x) - \nabla u(s,x)\| \leq C|t-s|^{\varepsilon/2}\|\nabla u\|^{1-1/q-\varepsilon/2}_{H^q_{2,p}}\|\partial_t u\|^{1/q+\varepsilon/2}_{L^q_p}, \tag{5.12}$$

$$\|\nabla u(t,x)\| + \frac{\|\nabla u(t,x) - \nabla u(t,y)\|}{|x-y|^\varepsilon} \leq CT^{-1/q}\left(\|u\|_{H^q_{2,p}} + T\|\partial_t u\|_{L^q_p}\right), \tag{5.13}$$

where $\|\cdot\|$ denotes any norm in $\mathbb{R}^{d\times d}$

Our method to construct strong solutions is actually motivated by the following observation in [74] and [82] (see also [84]).

Proposition 5.4. *Suppose that the drift coefficient $b : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ in (5.5) is bounded and Lipschitz continuous. Then the unique strong solution $X_t = (X^1_t, ..., X^d_t)$ of (5.5) has the explicit representation*

$$\varphi\left(t, X^i_t(\omega)\right) = E_{\widetilde{\mu}}\left[\varphi\left(t, \widetilde{B}^i_t(\widetilde{\omega})\right)\mathcal{E}^\diamond_T(b)\right] \tag{5.14}$$

for all $\varphi : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ such that $\varphi(t, B^i_t) \in L^2(\Omega)$ for all $0 \leq t \leq T$, $i = 1, \ldots, d$,. The random element $\mathcal{E}^\diamond_T(b)$ is given by

$$\begin{aligned}\mathcal{E}^\diamond_T(b)(\omega,\widetilde{\omega}) := \exp^\diamond\Bigg(&\sum_{j=1}^d \int_0^T \left(W^j_s(\omega) + b^j(s, \widetilde{B}_s(\widetilde{\omega}))\right) d\widetilde{B}^j_s(\widetilde{\omega}) \\ &- \frac{1}{2}\int_0^T \left(W^j_s(\omega) + b^j(s, \widetilde{B}_s(\widetilde{\omega}))\right)^{\diamond 2} ds\Bigg).\end{aligned} \tag{5.15}$$

Here $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mu}\right), \left(\widetilde{B}_t\right)_{t\geq 0}$ is a copy of the quadruple $(\Omega, \mathcal{F}, \mu), (B_t)_{t\geq 0}$ in (5.6). Further $E_{\widetilde{\mu}}$ denotes a Pettis integral of random elements $\Phi : \widetilde{\Omega} \longrightarrow (\mathcal{S})^*$ with respect to the measure

$\widetilde{\mu}$. The Wick product $\diamond$ in the Wick exponential of (5.15) (see 5.60) is taken with respect to μ and W_t^j is the white noise of B_t^j in the Hida space $(\mathcal{S})^*$ (see (5.57)). The stochastic integrals $\int_0^T \varphi(t,\widetilde{\omega})d\widetilde{B}_s^j(\widetilde{\omega})$ in (5.15) are defined for predictable integrands φ with values in the conuclear space $(\mathcal{S})^*$. See e.g. [61] for definitions. The other integral type in (5.15) is to be understood in the sense of Pettis.

Remark 5.5. *Let $0 = t_1^n < t_2^n < \ldots < t_{m_n}^n = T$ be a sequence of partitions of the interval $[0,T]$ with $\max_{i=1}^{m_n-1} \left|t_{i+1}^n - t_i^n\right| \longrightarrow 0$. Then the stochastic integral of the white noise W^j can be approximated as follows:*

$$\int_0^T W_s^j(\omega)d\widetilde{B}_s^j(\widetilde{\omega}) = \lim_{n\longrightarrow\infty}\sum_{i=1}^{m_n}(\widetilde{B}^j_{t^n_{i+1}}(\widetilde{\omega}) - \widetilde{B}^j_{t^n_i}(\widetilde{\omega}))W^j_{t^n_i}(\omega)$$

in $L^2(\lambda \times \widetilde{\mu}; (\mathcal{S})^*)$. For more information about stochastic integration on conuclear spaces the reader is referred to [61].

In the sequel we shall use the notation $Y_t^{i,b}$ for the expectation on the right hand side of (5.14) for $\varphi(t,x) = x$, that is

$$Y_t^{i,b} := E_{\widetilde{\mu}}\left[\widetilde{B}_t^{(i)}\mathcal{E}_T^{\diamond}(b)\right]$$

for $i = 1,\ldots,d$. We set

$$Y_t^b = \left(Y_t^{1,b},\ldots,Y_t^{d,b}\right) . \tag{5.16}$$

The form of Formula (5.14) in Proposition 5.4 actually gives rise to the conjecture that the expectation on the right hand side of Y_t^b in (5.16) may also define solutions of (5.5) for drift coefficients b lying in L_p^q.

Our method to construct strong solutions to SDE (5.5) which are Malliavin differentiable is essentially based on three steps.

- First, we consider a sequence of compactly supported smooth functions $b_n : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$, $n \geq 0$ such that $b_0 := b$ and $b_n \to b$ as $n \to \infty$ in L_p^q, then we prove that the sequence of strong solutions $X_t^n = Y_t^{b_n}$, $n \geq 1$, is relatively compact in $L^2(\Omega;\mathbb{R}^d)$ (Corollary 5.9) for every $t \in [0,T]$. The main tool to verify compactness is the bound in Lemma 5.6 in connection with a compactness criterion in terms of Malliavin derivatives obtained in [30] (see Appendix 5.B). This step is one of the main contributions of this paper.

- Secondly, given a merely measurable drift coefficient b in the space L_p^q, we show that Y_t^b, $t \in [0,T]$ is a generalized process in the Hida distribution space and we invoke the S-transform (5.58) to prove that for a given sequence of a.e. approximating, smooth coefficients b_n with compact support such that $\sup_{n\geq 0}\|b_n\|_{L_p^q}$, a subsequence of the corresponding strong solutions $X_t^{n_j} = Y_t^{b_{n_j}}$ fulfils

  $$Y_t^{b_{n_j}} \to Y_t^b$$

  in $L^2(\Omega;\mathbb{R}^d)$ for $0 \leq t \leq T$ (Lemma 5.12).

- Finally, using a certain transformation property for Y_t^b (Lemma 5.14) we directly show that Y_t^b is a Maaliavin differentiable solution to (5.5).

We turn now to the first step of our procedure. The successful completion of the first step relies on the following essential lemma:

Lemma 5.6. *Let $b_n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $n \geq 1$ be a sequence of functions in $C_0^\infty(\mathbb{R}^d)$ (space of infinitely often differentiable functions with compact support) approximating $b \in L_p^q$ a.e. such that $b_0 := b$ and $\sup_{n\geq 0} \|b_n\|_{L_p^q} < \infty$. Denote by $X_t^{n,x}$ the strong solution of SDE (5.5) with drift coefficient b_n for each $n \geq 0$. Then for every $t \in [0,T]$, $0 \leq r' \leq r \leq t$ there exist a $0 < \delta < 1$ and a continuous function $C : [0,\infty) \to [0,\infty)$ depending only on p, q, d, δ and T such that*

$$E\left[\|D_{r'}X_t^{n,x} - D_r X_t^{n,x}\|^2\right] \leq C(\|b_n\|_{L_p^q})|r' - r|^\delta \tag{5.17}$$

with

$$\sup_{n\geq 1} C(\|b_n\|_{L_p^q}) < \infty.$$

Here $\|\cdot\|$ denotes any norm in $\mathbb{R}^{d\times d}$.

Moreover,

$$\sup_{n\geq 1}\sup_{r\in[0,T]} E\left[\|D_r X_t^{n,x}\|^p\right] < \infty \tag{5.18}$$

for all $p \geq 2$.

Proof. Throughout the proof we will denote by $C_* : \mathbb{R} \to [0,\infty)$ any function depending on the parameters $*$. We will also use the symbol $\lesssim$ to denote *less or equal* up to a positive real constant independent of n.

We will prove the above estimates by considering the solution of the associated PDE presented in (5.9) with b_n, $n \geq 0$ in place of b which we denote by U_n, $n \geq 0$ and then using the results introduced at the beginning of this section on the regularity of its solution.

First, let us introduce a new process that will be useful for this purpose. Consider for each $n \geq 0$ and $t \in [0,T]$ the functions $\gamma_{t,n} : \mathbb{R}^d \to \mathbb{R}^d$ defined as $\gamma_{t,n}(x) = x + U_n(t,x)$. It turns out, see [44, Lemma 3.5], that the functions $\gamma_{t,n}$, $t \in [0,T]$, $n \geq 0$ define a family of C^1-diffeomorphisms on $\mathbb{R}^d$. Furthermore, consider the auxiliary process $\tilde{X}_t^{n,x} := \gamma_{t,n}(X_t^{n,x})$, $t \in [0,T]$, $n \geq 1$. One checks using Itô's formula and (5.9) that $\tilde{X}_t^{n,x}$ satisfies the following SDE

$$d\tilde{X}_t^{n,x} = \lambda U_n(t, \gamma_{t,n}^{-1}(\tilde{X}_t^{n,x}))dt + \left(\mathcal{I}_d + \nabla U_n(t, \gamma_{t,n}^{-1}(\tilde{X}_t^{n,x}))\right) dB_t, \quad \tilde{X}_0^{n,x} = x + U_n(0,x) \tag{5.19}$$

which is equivalent to SDE (5.5) if we replace b by b_n, $n \geq 1$. Using the chain rule for Malliavin derivatives (see e.g. [90]) we see that for $0 \leq r \leq t$,

$$D_r\tilde{X}_t^{n,x} = \nabla\gamma_{t,n}(X_t^{n,x})D_r X_t^{n,x}.$$

Because of Lemma 5.20 it suffices to prove the estimates (5.17) and (5.18) for the process $\tilde{X}_t^{n,x}$.

Since b_n are now smooth we have that (5.19) admits a unique strong solution which takes the form

$$\tilde{X}_t^{n,x} = x + U_n(0,x) + \lambda \int_0^t U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}))ds + \int_0^t \left(\mathcal{I}_d + \nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}))\right) dB_s.$$

Then the Malliavin derivative of $\tilde{X}_t^{n,x}$ for $0 \leq r \leq t$, which exists (see e.g. [90]), is

$$\begin{aligned} D_r \tilde{X}_t^{n,x} &= \mathcal{I}_d + \nabla U_n(r, \gamma_{r,n}^{-1}(\tilde{X}_r^{n,x})) \\ &\quad + \lambda \int_r^t \nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) D_r \tilde{X}_s^{n,x} ds \\ &\quad + \int_r^t \nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) D_r \tilde{X}_s^{n,x} dB_s. \end{aligned}$$

Denote for simplicity, $Z_{r,t}^n := D_r \tilde{X}_t^{n,x}$. Then for $r' < r$ we can write

$$\begin{aligned} Z_{r',t}^n - Z_{r,t}^n &= \nabla U_n(r', \gamma_{r',n}^{-1}(\tilde{X}_{r'}^{n,x})) - \nabla U_n(r, \gamma_{r,n}^{-1}(\tilde{X}_r^{n,x})) \\ &\quad + \lambda \int_{r'}^r \nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) Z_{r',s}^n ds \\ &\quad + \lambda \int_r^t \nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) \left(Z_{r',s}^n - Z_{r,s}^n\right) ds \\ &\quad + \int_{r'}^r \nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) Z_{r',s}^n dB_s \\ &\quad + \int_r^t \nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) \left(Z_{r',s}^n - Z_{r,s}^n\right) dB_s \\ &= Z_{r',r}^n - Z_{r,r}^n \\ &\quad + \lambda \int_r^t \nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) \left(Z_{r',s}^n - Z_{r,s}^n\right) ds \\ &\quad + \int_r^t \nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}) \left(Z_{r',s}^n - Z_{r,s}^n\right) dB_s. \end{aligned}$$

By dint of Lemma 5.19 we know that ∇U_n is bounded uniformly in n and Lemma 5.18 shows that $\nabla^2 U_n$ belongs, at least, to L_p^q uniformly in n. This implies that the stochastic integral in the expression for $Z_{r',t}^n - Z_{r,t}^n$ is a *true* martingale, which we here denote my M_t^n. As a result, since the initial condition $Z_{r',r}^n - Z_{r,r}^n$ is $\mathcal{F}_r$-measurable for each $n \geq 0$, for a given $\alpha \geq 2$, by Itô's formula we have

$$\begin{aligned} \|Z_{r',t}^n - Z_{r,t}^n\|^\alpha &\lesssim \|Z_{r',r}^n - Z_{r,r}^n\|^\alpha + \int_r^t \|Z_{r',s}^n - Z_{r,s}^n\|^\alpha ds + M_t^n \\ &\quad + \int_r^t \|Z_{r',s}^n - Z_{r,s}^n\|^{\alpha-2} \operatorname{tr}\left[\left(\nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})(Z_{r',s}^n - Z_{r,s}^n)\right)\right. \\ &\quad \left. \times \left(\nabla^2 U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})) \nabla \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})(Z_{r',s}^n - Z_{r,s}^n)\right)^* \right] ds \end{aligned} \tag{5.20}$$

where here tr stands for the trace and $*$ for the transposition of matrices.

We proceed using the fact that the trace of the matrix appearing in (5.20) can be bounded by a constant $C_{p,d}$ independent of n, times $\|Z^n_{r',s} - Z^n_{r,s}\|^2 \|\nabla^2 U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s))\nabla\gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)\|^2$.

Altogether,

$$\begin{aligned}\|Z^n_{r',t} - Z^n_{r,t}\|^\alpha \lesssim \|Z^n_{r',r} - Z^n_{r,r}\|^\alpha + \int_r^t \|Z^n_{r',s} - Z^n_{r,s}\|^\alpha ds + M^n_t \\ + \int_r^t \|Z^n_{r',s} - Z^n_{r,s}\|^\alpha \|\nabla^2 U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s))\nabla\gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)\|^2 ds\end{aligned} \tag{5.21}$$

Consider thus the process

$$V^n_t := \int_r^t \|\nabla^2 U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s))\nabla\gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)\|^2 ds. \tag{5.22}$$

The process V^n_t is a continuous non-decreasing and $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted process such that $V^n_r = 0$. Then Lemma 5.18 in connection with Theorem 5.2 we have that $\sup\limits_{n\geq 0} E[V^n_t] < \infty$.

Then Itô's formula yields

$$e^{-V^n_t}\|Z^n_{r',t} - Z^n_{r,t}\|^\alpha \lesssim \|Z^n_{r',r} - Z^n_{r,r}\|^\alpha + \int_r^t e^{-V^n_s}\|Z^n_{r',s} - Z^n_{r,s}\|^\alpha ds + \int_r^t e^{-V^n_s} dM_s. \tag{5.23}$$

Then taking expectation

$$E\left[e^{-V^n_t}\|Z^n_{r',t} - Z^n_{r,t}\|^\alpha\right] \lesssim E\left[\|Z^n_{r',r} - Z^n_{r,r}\|^\alpha\right] + \int_r^t E\left[e^{-V^n_s}\|Z^n_{r',s} - Z^n_{r,s}\|^\alpha\right] ds. \tag{5.24}$$

Then Grönwall's inequality gives

$$E\left[e^{-V^n_t}\|Z^n_{r',t} - Z^n_{r,t}\|^\alpha\right] \lesssim E\left[\|Z^n_{r',r} - Z^n_{r,r}\|^\alpha\right]. \tag{5.25}$$

At this point, it is easy to see, following similar steps, that for the process $Z^n_{r,t}$ one has

$$E\left[e^{-V^n_t}\|Z^n_{r,t}\|^\alpha\right] \lesssim E\left[\|Z^n_{r,r}\|^\alpha\right],$$

where $Z^n_{r,r} = \mathcal{I}_d + \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))$. So

$$\sup_{n\geq 0}\sup_{r\in[0,T]} E\left[e^{-V^n_t}\|Z^n_{r,t}\|^\alpha\right] \lesssim 1 + \sup_{n\geq 0}\sup_{r\in[0,T]} E\left[\|\nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))\|^\alpha\right] < \infty \tag{5.26}$$

because of Lemma 5.19 (ii) for a sufficiently large $\lambda \in \mathbb{R}$.

Then, the Cauchy-Schwarz inequality and Lemma 5.21 give

$$\sup_{n\geq 0}\sup_{r\in[0,T]} E\left[\|Z^n_{r,t}\|^\alpha\right] \leq \sup_{n\geq 0}\sup_{r\in[0,T]} E\left[e^{-2V^n_t}\|Z^n_{r,t}\|^{2\alpha}\right]^{1/2} \sup_{n\geq 0} E\left[e^{2V^n_T}\right]^{1/2} < \infty.$$

We continue to prove the estimate (5.17). Recall that

$$\begin{aligned} Z^n_{r',r} - Z^n_{r,r} &= \nabla U_n(r', \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r)) \\ &\quad + \lambda \int_{r'}^r \nabla U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)) \nabla \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s) Z_{r',s} ds \\ &\quad + \int_{r'}^r \nabla^2 U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)) \nabla \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s) Z_{r',s} dB_s. \end{aligned} \tag{5.27}$$

Then taking norm and using Burkholder-Davis-Gundy inequality we get

$$\begin{aligned} E\left[\|Z^n_{r',r} - Z^n_{r,r}\|^\alpha\right] &\lesssim E\left[\|\nabla U_n(r', \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))\|^\alpha\right] \\ &\quad + \lambda^\alpha E\left[\left(\int_{r'}^r \|\nabla U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)) \nabla \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s) Z_{r',s}\| ds\right)^\alpha\right] \\ &\quad + E\left[\left(\int_{r'}^r \|\nabla^2 U_n(s, \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s)) \nabla \gamma^{-1}_{s,n}(\tilde{X}^{n,x}_s) Z_{r',s}\|^2 ds\right)^{\alpha/2}\right]. \\ &=: i)_n + ii)_n + iii)_n \end{aligned} \tag{5.28}$$

The aim now is to find Hölder bounds in the sense of (5.17) for the expressions appearing in (5.28).

For $i)_n$ we may write

$$\begin{aligned} i)_n &= E\left[\|\nabla U_n(r', \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))\|^\alpha\right] \\ &\lesssim E\left[\|\nabla U_n(r', \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'}))\|^\alpha\right] \\ &\quad + E\left[\|\nabla U_n(r, \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))\|^\alpha\right]. \end{aligned}$$

Then by Lemma 5.3 there exists an $\varepsilon \in (0, 1/\alpha)$ and a constant $C_{p,q,d,\alpha} > 0$ independent of $n \geq 0$ such that

$$\begin{aligned} E&\left[\|\nabla U_n(r', \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'}))\|^\alpha\right] \\ &\qquad \leq C_{p,q,d,\alpha}\left(|r'-r|^{\varepsilon/2} \|\nabla U_n\|^{1-1/q-\varepsilon/2}_{H^q_{2,p}} \|\partial_t U_n\|^{1/q+\varepsilon/2}_{L^q_p}\right)^\alpha \end{aligned}$$

and

$$\begin{aligned} E&\left[\|\nabla U_n(r, \gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'})) - \nabla U_n(r, \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r))\|^\alpha\right] \\ &\qquad \leq C_{p,q,d,\alpha} T^{-\alpha/q} E\left[|\gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'}) - \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r)|^{\alpha\varepsilon}\right] \left(\|U_n\|_{H^q_{2,p}} + T\|\partial_t U_n\|_{L^q_p}\right)^\alpha. \end{aligned}$$

The above bounds in connection with inequality (5.11) in Theorem 5.2 give

$$i)_n \leq C_{p,q,d,\alpha,T}(\|b_n\|_{L^q_p})\left(|r'-r|^{\alpha\varepsilon/2} + E\left[|\gamma^{-1}_{r',n}(\tilde{X}^{n,x}_{r'}) - \gamma^{-1}_{r,n}(\tilde{X}^{n,x}_r)|^{\alpha\varepsilon}\right]\right)$$

for some continuous function $C_{p,q,d,\alpha,T}(\cdot)$ and hence

$$\sup_{n\geq 0} C_{p,q,d,\alpha,T}(\|b_n\|_{L_p^q}) < \infty.$$

Moreover, using Girsanov's theorem, we obtain that

$$\begin{aligned}
E\left[|\gamma_{r',n}^{-1}(\tilde{X}_{r'}^{n,x}) - \gamma_{r,n}^{-1}(\tilde{X}_{r}^{n,x})|\right] &= E\left[|X_{r'}^{n,x} - X_{r}^{n,x}|\right] \\
&\lesssim E\left[\left|\int_{r'}^{r} b_n(s, x+B_s)ds\right| \mathcal{E}\left(\int_0^T b_n(u, x+B_u)dB_u\right)\right] + E\left[|B_{r'} - B_r|\right] \\
&\lesssim |r'-r|^{1/2} E\left[\int_{r'}^{r} |b_n(s, x+B_s)|^2 ds\right]^{1/2} + |r'-r|^{1/2} \\
&\lesssim |r'-r|^{1/2},
\end{aligned}$$

where we used, Cauchy-Schwarz inequality and both that

$$\sup_{n\geq 0} E\left[\mathcal{E}\left(\int_0^T b_n(u, x+B_u)dB_u\right)^2\right] < \infty$$

and

$$\sup_{n\geq 0} E\left[\int_{r'}^{r} |b_n(s, x+B_s)|^2 ds\right]^{1/2} < \infty,$$

see [68, Lemma 3.2] or Lemma 5.17.

By Jensen's inequality for concave functions and the previous estimate we have

$$E\left[|\gamma_{r',n}^{-1}(\tilde{X}_{r'}^{n,x}) - \gamma_{r,n}^{-1}(\tilde{X}_{r}^{n,x})|^{\alpha\varepsilon}\right] \leq E\left[|\gamma_{r',n}^{-1}(\tilde{X}_{r'}^{n,x}) - \gamma_{r,n}^{-1}(\tilde{X}_{r}^{n,x})|\right]^{\alpha\varepsilon} \lesssim |r'-r|^{\alpha\varepsilon/2}.$$

Altogether,

$$i)_n \leq C_{p,q,d,\alpha,T}(\|b_n\|_{L_p^q})|r'-r|^{\delta}$$

for a $\delta \in (0,1)$.

For the second term, $ii)_n$, we use Hölder's inequality, Lemma 5.19 (ii) for a sufficiently large $\lambda \in \mathbb{R}$, Lemma 5.20 and the estimate (5.18) to obtain

$$\begin{aligned}
ii)_n &\lesssim \lambda^{\alpha}|r'-r|^{\alpha-1}\left(\sup_{s\in[0,t]} E\left[\|\nabla U_n(s, \gamma_{s,n}^{-1}(\tilde{X}_s^{n,x}))\nabla\gamma_{s,n}^{-1}(\tilde{X}_s^{n,x})\|^{2\alpha}\right] ds\right)^{1/2} \\
&\quad \times \left(\int_{r'}^{r} E\left[\|Z_{r',s}^{n}\|^{2\alpha}\right] ds\right)^{1/2} \\
&\leq C_{p,q,d,\alpha,T}|r'-r|^{\delta}
\end{aligned}$$

for a $\delta \in (0,1)$.

Finally, for the third term, for $\alpha \geq 2$, we use Hölder's inequality to obtain

$$iii)_n \lesssim |r'-r|^{\frac{\alpha-2}{2}} E\left[\int_{r'}^{r} \|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{\alpha} \|\nabla\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{\alpha} \|Z_{r',s}^n\|^{\alpha} ds\right].$$

Then choose $\alpha = 2(1+\delta)$ with $\delta \in (0,1/4)$ and use Lemma 5.20 to get

$$iii)_n \lesssim |r'-r|^{\delta} E\left[\int_{r'}^{r} \|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)} \|Z_{r',s}^n\|^{2(1+\delta)} ds\right].$$

Then Fubini's theorem, Hölder's inequality once more with respect to $\mu(d\omega)$, with exponent $1+\delta'$, $\delta' \in (0,1/4)$ and Cauchy-Schwarz yield

$$\begin{aligned}
&E\left[\int_{r'}^{r} \|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)} \|Z_{r',s}^n\|^{2(1+\delta)} ds\right] \\
&\quad = \int_{r'}^{r} E\left[\|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)} \|Z_{r',s}^n\|^{2(1+\delta)}\right] ds \\
&\quad \lesssim \int_{r'}^{r} E\left[\|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)(1+\delta')}\right]^{1/(1+\delta')} E\left[\|Z_{r',s}^n\|^{2(1+\delta)\frac{1+\delta'}{\delta'}}\right]^{\frac{\delta'}{1+\delta'}} ds \\
&\quad \lesssim \sup_{n\geq 0}\sup_{s\in[r',r]} E\left[\|Z_{r',s}^n\|^{2(1+\delta)\frac{1+\delta'}{\delta'}}\right]^{\frac{\delta'}{1+\delta'}} \int_{r'}^{r} E\left[\|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)(1+\delta')}\right]^{1/(1+\delta')} ds \\
&\quad \lesssim \int_0^T E\left[\|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)(1+\delta')}\right]^{1/(1+\delta')} ds
\end{aligned}$$

where the last step follows from (5.18). For the last factor, since $0 < 1/(1+\delta') < 1$, using the inverse Jensen's inequality and the fact that $1 < (1+\delta)(1+\delta') < 2$ for suitable $\delta, \delta' \in (0,1/4)$ in connection with Lemma 5.18 we have

$$\begin{aligned}
&\int_0^T E\left[\|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)(1+\delta')}\right]^{1/(1+\delta')} ds \\
&\qquad \leq T^{1-1/(1+\delta')}\left(E\left[\int_0^T \|\nabla^2 U_n(s,\gamma_{s,n}^{-1}(\tilde{X}_s^n))\|^{2(1+\delta)(1+\delta')} ds\right]\right)^{1/(1+\delta')} \leq M < \infty
\end{aligned}$$

for every $n \geq 0$, w.r.t. a constant M.

As a summary, it follows from (5.25) that

$$E\left[e^{-V_t^n}\|Z_{r',t}^n - Z_{r,t}^n\|^{2(1+\delta)}\right] \leq C_{p,q,d,\alpha,T}(\|b_n\|_{L_p^q})|r'-r|^{\delta}.$$

Then by Hölder's inequality with exponent $1+\delta$, $\delta \in (0,1)$ together with Lemma 5.21 we obtain

$$\begin{aligned}
E\left[\|Z_{r',t}^n - Z_{r,t}^n\|^2\right] &= E\left[e^{\frac{1}{1+\delta}V_t^n} e^{-\frac{1}{1+\delta}V_t^n}\|Z_{r',t}^n - Z_{r,t}^n\|^2\right] \\
&\leq E\left[e^{\frac{1}{\delta}V_t^n}\right]^{\frac{\delta}{1+\delta}} E\left[e^{-V_t^n}\|Z_{r',t}^n - Z_{r,t}^n\|^{2(1+\delta)}\right]^{\frac{1}{1+\delta}} \\
&\leq C_{p,q,d,\alpha,T}(\|b_n\|_{L_p^q})|r'-r|^{\delta/(1+\delta)}
\end{aligned}$$

for some continuous function $C : [0,\infty) \to [0,\infty)$ such that

$$\sup_{n\geq 0} C_{p,q,d,\alpha,T}(\|b_n\|_{L_p^q}) < \infty.$$

□

Remark 5.7. *The bound given in (5.18) is in fact uniform in $x \in \mathbb{R}^d$. Indeed, by Lemma 5.19 item (ii) we have that the bound given in (5.26) is also uniform in $x \in \mathbb{R}^d$. Moreover, since $\Delta U_n \in L_p^q$ for all $n \geq 0$, then by Lemma 5.19 item (iii) in connection with Lemma 5.17 we have that for any $k \in \mathbb{R}$*

$$\sup_{x\in\mathbb{R}^d} \sup_{n\geq 0} E[e^{kV_T^n}] < \infty.$$

Hence, for any $\alpha \geq 1$

$$\sup_{x\in\mathbb{R}^d} \sup_{r\in[0,T]} \sup_{n\geq 0} E\left[\|D_r X_t^{n,x}\|^\alpha\right] < \infty.$$

Remark 5.8. *One also checks that the same holds for the spatial derivatives, that is for any $\alpha \geq 1$*

$$\sup_{x\in\mathbb{R}^d} \sup_{r\in[0,T]} \sup_{n\geq 0} E\left[\left\|\frac{\partial}{\partial x} X_t^{n,x}\right\|^\alpha\right] < \infty$$

by using the fact that $\frac{\partial}{\partial x} X_t^{n,x}$ solves the same SDE as $D_r X_t^{n,x}$, starting at $r = 0$.

As a repercussion of Lemma 5.6 we have the following result which is central in the proof of the existence of strong solutions of (5.5).

Corollary 5.9. *Let $\{b_n\}_{n\geq 0}$ be a sequence of compactly supported smooth functions approximating b in L_p^q. Denote, as before, $X_t^{x,n}$ the solution to equation (5.5) with drift coefficient b_n. Then for each $t \in [0,T]$ the sequence of random variables $X_t^{n,x}$, $n \geq 0$ is relatively compact in $L^2(\Omega)$.*

Proof. This is a direct consequence of the compactness criterion that can be found in Appendix 5.C, Lemma 5.22 and 5.23, which is due to [30], together with Lemma 5.6. One can check that the double integral in Lemma 5.23 is finite. Namely

$$\int_0^T \int_0^T \frac{E\left[\|Z_{r',t}^n - Z_{r,t}\|^2\right]}{|r'-r|^{1+2\beta}} dr' dr \leq \int_0^T \int_0^T \frac{1}{|r'-r|^{2\beta+1-\delta}} dr' dr < \infty$$

for any $0 < \delta < 1$ and $2\beta + 1 - \delta < 1$. □

The following lemma gives a criterion under which the process Y_t^b belongs to the Hida distribution space.

Lemma 5.10. *Suppose that*

$$E_\mu\left[\exp\left(36 \int_0^T |b(s,B_s)|^2 ds\right)\right] < \infty, \tag{5.29}$$

where the drift $b : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is measurable (in particular, (5.29) is valid for $b \in L_p^q$ because of Lemma (5.17)). Then the coordinates of the process Y_t^b, defined in (5.16), that is

$$Y_t^{i,b} = E_{\widetilde{\mu}}\left[\widetilde{B}_t^{(i)}\mathcal{E}_T^{\diamond}(b)\right], \tag{5.30}$$

are elements of the Hida distribution space.

Proof. See [83] for a similar proof. □

Lemma 5.11. *Let $\varepsilon \in (0,1)$ and define $p_\varepsilon := 1+\varepsilon$ and $q_\varepsilon := \frac{1+\varepsilon}{\varepsilon}$. Let $b_n : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be a sequence of Borel measurable functions with $b_0 = b$ such that*

$$\sup_{n\geq 0} E\left[\exp\left(16q_\varepsilon(8q_\varepsilon - 1)\int_0^T |b_n(s,B_s)|^2\, ds\right)\right] < \infty \tag{5.31}$$

holds. Then

$$\left|S(Y_t^{i,b_n} - Y_t^{i,b})(\varphi)\right| \leq const \cdot E[J_n]^{\frac{1}{p_\varepsilon}} \cdot \exp\left(2(8q_\varepsilon - 1)\int_0^T |\varphi(s)|^2\, ds\right)$$

for all $\varphi \in (S_{\mathbb{C}}([0,T]))^d$, $i = 1,\dots,d$, where S denotes the S-transform (see Section 5.A.1 in Appendix 5.A) and where the factor J_n is defined by

$$J_n = \sum_{j=1}^d 2\left|\int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds\right|^{\frac{p_\varepsilon}{2}} + \left|\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2) ds\right|^{p_\varepsilon}. \tag{5.32}$$

Here $S_{\mathbb{C}}([0,T])$ is the complexification of the Schwarz space $S([0,T])$ on $[0,T]$, see Section 5.A.1 in Appendix 5.A.

In particular, if b_n approximates b in the following sense

$$E[J_n] \to 0 \tag{5.33}$$

as $n \to \infty$, it follows that

$$Y_t^{b_n} \to Y_t^b \text{ in } (\mathcal{S})^*$$

as $n \to \infty$ for all $0 \leq t \leq T$, $i = 1,\dots,d$.

Proof. For $i = 1,\dots,d$ we obtain by Proposition 5.4 and (5.59) that

$$\begin{aligned}|S(Y_t^{i,b_n} - Y_t^{i,b})(\varphi)| \leq E_{\tilde{\mu}}\Bigg[|\tilde{B}_t^{(i)}| \exp\Bigg\{\sum_{j=1}^d \mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)} \\ - \frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\} \\ \times \Bigg|\exp\Bigg\{\sum_{j=1}^d \int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))d\tilde{B}_s^{(j)}\end{aligned}$$

$$+\frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2)ds$$
$$\left. + \int_0^T \varphi^{(j)}(s)(b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))ds \right\} - 1 \Bigg| \Bigg].$$

Since $|\exp\{z\} - 1| \leq |z|\exp\{|z|\}$ it follows from Hölder's inequality with exponents $p_\varepsilon = 1+\varepsilon$ and $q_\varepsilon = \frac{1+\varepsilon}{\varepsilon}$, for an appropriate $\varepsilon > 0$, that

$$|S(Y_t^{i,b_n} - Y_t^{i,b})(\varphi)| \leq E_{\tilde{\mu}}\left[|Q_n|^{p_\varepsilon}\right]^{\frac{1}{p_\varepsilon}} E_{\tilde{\mu}}\Bigg[\Bigg(|\tilde{B}_t^{(i)}|\exp\Bigg\{\sum_{j=1}^d \mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})$$
$$+ \varphi^{(j)}(s))d\tilde{B}_s^{(j)} - \frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg)^{q_\varepsilon} \exp\{q_\varepsilon |Q_n|\}\Bigg]^{\frac{1}{q_\varepsilon}},$$

where

$$Q_n = \sum_{j=1}^d \int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))d\tilde{B}_s^{(j)} + \frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2)ds$$
$$+ \int_0^T \varphi^{(j)}(s)(b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))ds.$$

Then using the Cauchy-Schwarz inequality on the last integral and the fact that $|x| \leq e^x$ and $1 \leq e^x$ for $x \geq 0$ we may write

$$E_{\tilde{\mu}}\left[|Q_n|^{p_\varepsilon}\right] \leq C\exp\left\{\left(\int_0^T |\varphi(s)|^2 ds\right)^{p_\varepsilon/2}\right\} E_{\tilde{\mu}}\Bigg[\sum_{j=1}^d \left|\int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))d\tilde{B}_s^{(j)}\right|^{p_\varepsilon}$$
$$+ \left|\frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2)ds\right|^{p_\varepsilon}$$
$$+ \left|\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds\right|^{\frac{p_\varepsilon}{2}}\Bigg]$$
$$= C\exp\left\{\left(\int_0^T |\varphi(s)|^2 ds\right)^{p_\varepsilon/2}\right\} E_{\tilde{\mu}}\Bigg[\sum_{j=1}^d 2\left|\int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds\right|^{\frac{p_\varepsilon}{2}}$$
$$+ \left|\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2)ds\right|^{p_\varepsilon}\Bigg],$$

where in the last inequality we used the Burkholder-Davis-Gundy inequality for the stochastic integral. Then

$$E_{\tilde{\mu}}\left[|Q_n|^{p_\varepsilon}\right]^{\frac{1}{p_\varepsilon}} \leq C\exp\left\{\frac{1}{p_\varepsilon}\left(\int_0^T |\varphi(s)|^2 ds\right)^{p_\varepsilon/2}\right\} E_{\tilde{\mu}}\left[J_n\right]^{\frac{1}{p_\varepsilon}},$$

where

$$J_n = \sum_{j=1}^{d} 2\left|\int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds\right|^{\frac{p_\varepsilon}{2}} + \left|\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2) ds\right|^{p_\varepsilon}.$$

Further we get that

$$\begin{aligned}
&E_{\tilde{\mu}}\Bigg[\Bigg(|\tilde{B}_t^{(i)}|\exp\Bigg\{\sum_{j=1}^{d}\mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)} \\
&\qquad - \frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg)^{q_\varepsilon} \exp\{q_\varepsilon |Q_n|\}\Bigg]^{\frac{1}{q_\varepsilon}} \\
&\quad \le E_{\tilde{\mu}}\Bigg[\Bigg(|\tilde{B}_t^{(i)}|\exp\Bigg\{\sum_{j=1}^{d}\mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)} \\
&\qquad - \frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg)^{2q_\varepsilon}\Bigg]^{\frac{1}{2q_\varepsilon}} E_{\tilde{\mu}}\Bigg[\exp\{2q_\varepsilon |Q_n|\}\Bigg]^{\frac{1}{2q_\varepsilon}}.
\end{aligned}$$

Then for $z \in \mathbb{C}$ one has $\exp\{|z|\} \le \frac{1}{2}\Big(\exp\{2\mathrm{Re}\, z\} + \exp\{-2\mathrm{Re}\, z\} + \exp\{2\mathrm{Im}\, z\} + \exp\{-2\mathrm{Im}\, z\}\Big)$. Thus

$$\begin{aligned}
E_{\tilde{\mu}}\Bigg[\exp\{2q_\varepsilon|Q_n|\}\Bigg]^{\frac{1}{2q_\varepsilon}} &\le \frac{1}{2^{2q_\varepsilon}}\Bigg(E_{\tilde{\mu}}\Bigg[\exp\{4q_\varepsilon \mathrm{Re}\, Q_n\}\Bigg]^{\frac{1}{2q_\varepsilon}} + E_{\tilde{\mu}}\Bigg[\exp\{-4q_\varepsilon \mathrm{Re}\, Q_n\}\Bigg]^{\frac{1}{2q_\varepsilon}} \\
&\qquad + E_{\tilde{\mu}}\Bigg[\exp\{4q_\varepsilon \mathrm{Im}\, Q_n\}\Bigg]^{\frac{1}{2q_\varepsilon}} + E_{\tilde{\mu}}\Bigg[\exp\{-4q_\varepsilon \mathrm{Im}\, Q_n\}\Bigg]^{\frac{1}{2q_\varepsilon}}\Bigg).
\end{aligned}$$

By the Cauchy-Schwarz inequality and the supermartingale property of Doléans-Dade exponentials we get

$$\begin{aligned}
E_{\tilde{\mu}}\Bigg[\exp\{4q_\varepsilon \mathrm{Re}\, Q_n\}\Bigg] &\le E_{\tilde{\mu}}\Bigg[\exp\Bigg\{\sum_{j=1}^{d} 32q_\varepsilon^2 \int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds \\
&\qquad + 4q_\varepsilon \int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2) ds \\
&\qquad + 8q_\varepsilon \int_0^T \mathrm{Re}\, \varphi^{(j)}(s)(b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))ds\Bigg]\Bigg\}\Bigg]^{\frac{1}{2}} \\
&\le L_n \exp\left\{2q_\varepsilon \int_0^T |\varphi(s)|^2 ds\right\},
\end{aligned}$$

where the last step follows from the fact that $\langle f, g\rangle \le \frac{1}{2}(\|f\|^2 + \|g\|^2)$, $f, g \in L^2([0,T])$ and

where

$$L_n = E_{\tilde{\mu}}\Bigg[\exp\Bigg\{\sum_{j=1}^{d} 4q_\varepsilon(8q_\varepsilon+1)\int_0^T (b_n^{(j)}(s,\tilde{B}_s^{(j)}) - b^{(j)}(s,\tilde{B}_s^{(j)}))^2 ds$$

$$+ 4q_\varepsilon \int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2) ds\Bigg]\Bigg\}\Bigg]^{\frac{1}{2}}.$$

Similarly, one also obtains

$$E_{\tilde{\mu}}\left[\exp\{-4q_\varepsilon \mathrm{Re}\, Q_n\}\right] \leq L_n \exp\left\{2q_\varepsilon \int_0^T |\varphi(s)|^2 ds\right\}.$$

In the same way, one also obtains the same bounds for both $E_{\tilde{\mu}}\left[\exp\{4q_\varepsilon \mathrm{Im}\, Q_n\}\right]$ and $E_{\tilde{\mu}}\left[\exp\{-4q_\varepsilon \mathrm{Im}\, Q_n\}\right]$.

Finally, for the remaining factor we see that

$$E_{\tilde{\mu}}\Bigg[\Bigg(|\tilde{B}_t^{(i)}|\exp\Bigg\{\sum_{j=1}^{d}\mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)}$$

$$-\frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg)^{2q_\varepsilon}\Bigg]^{\frac{1}{2q_\varepsilon}}$$

$$\leq E_{\tilde{\mu}}\left[|\tilde{B}_t^{(i)}|^{4q_\varepsilon}\right]^{\frac{1}{4q_\varepsilon}} E_{\tilde{\mu}}\Bigg[\exp\Bigg\{4q_\varepsilon\sum_{j=1}^{d}\mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)}$$

$$-\frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg]^{\frac{1}{4q_\varepsilon}}$$

$$\leq E_{\tilde{\mu}}\left[|\tilde{B}_t^{(i)}|^{4q_\varepsilon}\right]^{\frac{1}{4q_\varepsilon}} E_{\tilde{\mu}}\Bigg[\exp\Bigg\{\sum_{j=1}^{d} 4q_\varepsilon(8q_\varepsilon-1)\int_0^T \mathrm{Re}\,(b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg\}\Bigg]^{\frac{1}{4q_\varepsilon}}.$$

Now, since $\mathrm{Re}\,(z^2) \leq (\mathrm{Re}\, z)^2$, $z \in \mathbb{C}$ we have that $\mathrm{Re}\,(b+\varphi)^2 \leq (b + \mathrm{Re}\,\varphi)^2$ then using Minkowski's inequality, i.e. $\|f+g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p)$ for any $p \geq 1$ and Cauchy-Schwarz inequality w.r.t. $\tilde{\mu}$ one finally obtains

$$E_{\tilde{\mu}}\Bigg[\Bigg(|\tilde{B}_t^{(i)}|\exp\Bigg\{\sum_{j=1}^{d}\mathrm{Re}\Bigg[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))d\tilde{B}_s^{(j)}$$

$$-\frac{1}{2}\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + \varphi^{(j)}(s))^2 ds\Bigg]\Bigg\}\Bigg)^{2q_\varepsilon}\Bigg]^{\frac{1}{2q_\varepsilon}}$$

$$\leq CE_{\tilde{\mu}}\left[\exp\left\{16q_\varepsilon(8q_\varepsilon-1)\int_0^T |b(s,\tilde{B}_s)|^2 ds\right\}\right]^{\frac{1}{8q_\varepsilon}} \exp\left\{2(8q_\varepsilon-1)\int_0^T |\varphi(s)|^2 ds\right\}.$$

Altogether, we obtain

$$\left|S(Y_t^{i,b_n} - Y_t^{i,b})(\varphi)\right| \leq const \cdot E[J_n]^{\frac{1}{1+\varepsilon}} \cdot \exp\left\{2\left(8\frac{1+\varepsilon}{\varepsilon} - 1\right)\int_0^T |\varphi(s)|^2 ds\right\}.$$

□

Lemma 5.12. *Let $b_n : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be a sequence of smooth functions with compact support with $b_0 := b$ which approximate the coefficient $b : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ in L_p^q. Then for any $0 \leq t \leq T$ there exists a subsequence of the corresponding strong solutions $X_{n_j,t} = Y_t^{b_{n_j}}$, $j = 1, 2...$, such that*

$$Y_t^{b_{n_j}} \longrightarrow Y_t^b$$

for $j \to \infty$ in $L^2(\Omega)$. In particular this implies $Y_t^b \in L^2(\Omega)$, $0 \leq t \leq T$.

Proof. By Corollary 5.9 we know that there exists a subsequence $Y_t^{b_{n_j}}$, $j \geq 1$, converging in $L^2(\Omega)$. Further, we need to show that $E[J_{n_j}] \to 0$ as $j \to \infty$ with J_{n_j} as in (5.32). To this end, observe that for a function $f \in L_p^q$ one has

$$E\left[\int_0^T f(s, \tilde{B}_s) ds\right] = \int_0^T (2\pi s)^{-d/2} \int_{\mathbb{R}^d} f(s,z) e^{-|z|^2/(2s)} dz ds.$$

Then by using Hölder's inequality with respect to z and then to s we see that for any $p', q' \in [1, \infty]$ satisfying

$$\frac{d}{p'} + \frac{2}{q'} < 2,$$

we have

$$E\left[\int_0^T f(s, \tilde{B}_s) ds\right] \leq C \|f\|_{L_{p'}^{q'}},$$

where C is a constant depending on T, d, p', q'. Then from condition (5.7), since $p, q > 2$ we can find an $\delta \in [0, 1)$ small enough so that $p, q > 2(1+\delta)$. For these p, q define $p' := \frac{p}{2(1+\delta)} \geq 1$ and $q := \frac{q'}{2(1+\delta)} > 1$ and apply the above estimate to $|f|^{2(1+\delta)}$ to obtain

$$E\left[\int_0^T |f(s, \tilde{B}_s)|^{2(1+\delta)} ds\right] \leq C \|f\|_{L_p^q}. \tag{5.34}$$

Now since $b_n^{(j)} - b^{(j)} \in L_p^q$ for every $j = 1, \ldots, d$ and $0 < \frac{1+\varepsilon}{2} < 1$ we have

$$E\left[\left(\int_0^T (b_n^{(j)}(s, \tilde{B}_s^{(j)}) - b^{(j)}(s, \tilde{B}_s^{(j)}))^2 ds\right)^{\frac{1+\varepsilon}{2}}\right] \leq E\left[\int_0^T (b_n^{(j)}(s, \tilde{B}_s^{(j)}) - b^{(j)}(s, \tilde{B}_s^{(j)}))^2 ds\right]^{\frac{1+\varepsilon}{2}}$$

which goes to zero by the above estimate (5.34) by just taking the case where $\delta = 0$.

Finally, for the the second term in $E[J_{n_j}]$ we have

$$
\begin{aligned}
& E\left[\left|\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)})^2 - b_n^{(j)}(s,\tilde{B}_s^{(j)})^2)ds\right|^{1+\varepsilon}\right] \\
\leq & T^\varepsilon E\left[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{1+\varepsilon}(b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{1+\varepsilon} ds\right] \\
\leq & T^\varepsilon \int_0^T E\left[(b^{(j)}(s,\tilde{B}_s^{(j)}) + b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)}\right]^{1/2} E\left[(b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)}\right]^{1/2} ds \\
\leq & T^\varepsilon E\left[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)} ds\right]^{1/2} \\
& \times E\left[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)} ds\right]^{1/2}.
\end{aligned}
$$

Then since $b^{(j)} + b_n(j) \in L_p^q$ for every $n \geq 0$ we have

$$
\sup_{n\geq 0} E\left[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) + b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)} ds\right]^{1/2} < \infty
$$

for a sufficiently small $\varepsilon \in (0,1)$ by Lemma 5.18 and

$$
E\left[\int_0^T (b^{(j)}(s,\tilde{B}_s^{(j)}) - b_n^{(j)}(s,\tilde{B}_s^{(j)}))^{2(1+\varepsilon)} ds\right]^{1/2} \to 0
$$

as $n \to \infty$ by estimate (5.34) for a sufficiently small $\varepsilon > 0$.

Thus, by Lemma 5.11, $Y_t^{b_{n_j}} \to Y_t^b$ as $j \to \infty$ in $(\mathcal{S})^*$. But then, by uniqueness of the limit, also $Y_t^{b_{n_j}} \to Y_t^b$ in $L^2(\Omega)$. □

Remark 5.13. *It follows from the above proof that $Y_t^{b_n} \to Y_t^b$ as $n \to \infty$ in $L^2(\Omega;\mathbb{R}^d)$ for all t and x.*

In fact, Lemma 5.12 enables us now to state the following "transformation property" for Y_t^b.

Lemma 5.14. *Assume that $b : [0,T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is in L_p^q. Then*

$$
\varphi^{(i)}\left(t, Y_t^b\right) = E_{\widetilde{\mu}}\left[\varphi^{(i)}\left(t, \widetilde{B}_t\right)\mathcal{E}_T^\diamond(b)\right] \tag{5.35}
$$

a.e. for all $0 \leq t \leq T$, $i = 1,\ldots,d$ and $\varphi = (\varphi^{(1)},\ldots,\varphi^{(d)})$ such that $\varphi(B_t) \in L^2(\Omega;\mathbb{R}^d)$.

Proof. See [99, Lemma 16] or [82]. □

Using the above auxiliary results we can finally give the proof of Theorem 5.1.

Proof of Theorem 5.1. We want to use the transformation property (5.35) of Lemma 5.14 to show that Y_t^b is a unique strong solution of the SDE (5.5). In order to shorten notation we set $\int_0^t \varphi(s,\omega)dB_s := \sum_{j=1}^d \int_0^t \varphi^{(j)}(s,\omega)dB_s^{(j)}$ and $x = 0$. Also, let b_n, $n = 1, 2, \ldots$, be a sequence of functions as required in Lemma 5.12.

We comment on that $Y_{\cdot}^{b}$ has a continuous modification. The latter can be seen as follows: Since each $Y_t^{b_n}$ is a strong solution of the SDE (5.5) with respect to the drift b_n we obtain from Girsanov's theorem and our assumptions that

$$\begin{aligned} E_\mu\left[\left(Y_t^{i,b_n}-Y_u^{i,b_n}\right)^4\right] &= E_{\widetilde{\mu}}\left[\left(\widetilde{B}_t^{(i)}-\widetilde{B}_u^{(i)}\right)^4\mathcal{E}\left(\int_0^T b_n(s,\widetilde{B}_s)d\widetilde{B}_s\right)\right] \\ &\leq const\cdot|t-u|^2 \end{aligned}$$

for all $0\leq u,t\leq T$, $n\geq 1$, $i=1,...,d$. The above constant comes from the fact that $\left\{\mathcal{E}\left(\int_0^T b_n(s,\widetilde{B}_s)d\widetilde{B}_s\right)\right\}_{n\geq 1}$ is bounded in $L^2(\Omega;\mathbb{R}^d)$ with respect to the measure $\tilde{\mu}$, see Lemma 3.2. in [68] or Lemma 5.17.

By Remark 5.13 we know that

$$Y_t^{b_n}\longrightarrow Y_t^b \text{ in } L^2(\Omega;\mathbb{R}^d)$$

and hence we have almost sure convergence for a further subsequence, $0\leq t\leq T$. So we get that by Fatou's lemma

$$E_\mu\left[\left(Y_t^{i,b}-Y_u^{i,b}\right)^4\right]\leq const\cdot|t-u|^2 \tag{5.36}$$

for all $0\leq u,t\leq T$, $i=1,...,d$. Then Kolmogorov's lemma guarantees a continuous modification of Y_t^b.

Since $\widetilde{B}_t$ is a weak solution of (5.5) for the drift $b(s,x)+\varphi(s)$ with respect to the measure $d\mu^*=\mathcal{E}\left(\int_0^T\left(b(s,\widetilde{B}_s)+\varphi(s)\right)d\widetilde{B}_s\right)d\mu$ we get that

$$\begin{aligned} S(Y_t^{i,b})(\varphi) &= E_{\widetilde{\mu}}\left[\widetilde{B}_t^{(i)}\mathcal{E}\left(\int_0^T\left(b(s,\widetilde{B}_s)+\varphi(s)\right)d\widetilde{B}_s\right)\right] \\ &= E_{\mu^*}\left[\widetilde{B}_t^{(i)}\right] \\ &= E_{\mu^*}\left[\int_0^t\left(b^{(i)}(s,\widetilde{B}_s)+\varphi^{(i)}(s)\right)ds\right] \\ &= \int_0^t E_{\widetilde{\mu}}\left[b^{(i)}(s,\widetilde{B}_s)\mathcal{E}\left(\int_0^T\left(b(u,\widetilde{B}_u)+\varphi(u)\right)d\widetilde{B}_u\right)\right]ds+S\left(B_t^{(i)}\right)(\varphi). \end{aligned}$$

Thus the transformation property (5.35) applied to b yields

$$S(Y_t^{i,b})(\varphi)=S(\int_0^t b^{(i)}(u,Y_u^{i,b})du)(\varphi)+S(B_t^{(i)})(\varphi).$$

Then it follows from the injectivity of the S-transform that

$$Y_t^b=\int_0^t b(s,Y_s^b)ds+B_t\,.$$

See Section 5.A in the Appendix.

The Malliavin differentiability of Y_t^b comes from the fact that $Y_t^{i,b_n} \to Y_t^{i,b}$ in $L^2(\Omega)$ and

$$\sup_{n\geq 1} \|Y_t^{i,b_n}\|_{\mathbb{D}^{1,2}} \leq M < \infty$$

for all $i = 1, \ldots, d$ and $0 \leq t \leq 1$. See e.g. [90].

On the other hand, using uniqueness in law, which is a consequence of Lemma 5.18 and Proposition 3.10, Ch. 5 in [62] we may apply, under our conditions, Girsanov's theorem to any other solution. Then the proof of Proposition 5.4 (see e.g. [98, Proposition 1]) shows that any other solution necessarily coincides with Y_t^b. □

We conclude this section with a generalisation of Theorem 5.1 to a class of non-degenerate $d-$dimensional Itô-diffusions.

Theorem 5.15. *Assume the time-homogeneous $\mathbb{R}^d-$valued SDE*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \quad X_0 = x \in \mathbb{R}^d, \quad 0 \leq t \leq T, \tag{5.37}$$

where the coefficients $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^d \times \mathbb{R}^d$ are Borel measurable. Suppose that there exists a bijection $\Lambda : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, which is twice continuously differentiable. Let $\Lambda_x : \mathbb{R}^d \longrightarrow L\left(\mathbb{R}^d, \mathbb{R}^d\right)$ and $\Lambda_{xx} : \mathbb{R}^d \longrightarrow L\left(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d\right)$ be the corresponding derivatives of Λ and assume that

$$\Lambda_x(y)\sigma(y) = id_{\mathbb{R}^d} \text{ for } y \text{ a.e.}$$

as well as

$$\Lambda^{-1} \text{ is Lipschitz continuous.}$$

Require that the function $b_* : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ given by

$$\begin{aligned} b_*(x) &:= \Lambda_x\left(\Lambda^{-1}(x)\right)\left[b(\Lambda^{-1}(x))\right] \\ &+ \frac{1}{2}\Lambda_{xx}\left(\Lambda^{-1}(x)\right)\left[\sum_{i=1}^d \sigma(\Lambda^{-1}(x))\left[e_i\right], \sum_{i=1}^d \sigma(\Lambda^{-1}(x))\left[e_i\right]\right] \end{aligned}$$

satisfies the conditions of Theorem 5.1, where e_i, $i = 1, \ldots, d$, is a basis of $\mathbb{R}^d$. Then there exists a Malliavin differentiable solution X_t to (5.37).

Proof. The proof can be directly obtained from Itô's Lemma. See [83]. □

5.3 Applications

5.3.1 The Bismut-Elworthy-Li formula

As an application we want to use Theorem 5.1 to derive a Bismut-Elworthy-Li formula for solutions v to the Kolmogorov equation

$$\frac{\partial}{\partial t}v(t,x) = \sum_{j=1}^d b_j(t,x)\frac{\partial}{\partial x_j}v(t,x) + \frac{1}{2}\sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}v(t,x) \tag{5.38}$$

with initial condition $v(0,x) = \Phi(x)$, where $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ belongs to L_p^q.

It is known that, see [68] or [46], that when Φ is continuous and bounded there exists a solution to (5.38) given by

$$v(t,x) = E[\Phi(X_t^x)], \tag{5.39}$$

where v is a solution to the Kolmogorov Equation (5.38) which is unique among all bounded solutions in the space $H_{2,p}^q$, as introduced in Theorem 5.2, with $p,q > 2$ satisfying (5.7). Moreover, $\frac{\partial}{\partial x} v \in L^\infty([0,T] \times \mathbb{R}^d)$.

In the sequel, we aim at finding a representation for $\frac{\partial}{\partial x} v$ without using derivatives of Φ. See [81] in the case of $b \in L^\infty([0,T] \times \mathbb{R}^d)$.

Theorem 5.16 (Bismut-Elworthy-Li formula)**.** *Assume $\Phi \in C_b(\mathbb{R}^d)$ and let U be an open, bounded subset of $\mathbb{R}^d$. Then the derivative of the solution to (5.38) can be represented as*

$$\frac{\partial}{\partial x} v(t,x) = E[\Phi(X_t^x) \int_0^t a(s) \left(\frac{\partial}{\partial x} X_s^x \right)^* dB_s]^* \tag{5.40}$$

for almost all $x \in U$ and all $t \in (0,T]$, where $a = a_t$ is any bounded measurable function such that $\int_0^t a_t(s) ds = 1$ and where $*$ denotes the transposition of matrices.

Proof. The proof is similar to Theorem 2 in [83] in the case of $b \in L^\infty([0,T] \times \mathbb{R}^d)$. For the convenience of the reader we give the full proof.

Assume that $\Phi \in C_b^2(\mathbb{R}^d)$ (the general case of $\Phi \in C_b(\mathbb{R}^d)$ can be proved by approximation of Φ in relation (5.42)) and let b_n and $X_t^{n,x}$ be as in the previous section. If we replace b by b_n in (5.38) we have the unique solution given by

$$v_n(t,x) = E[\Phi(X_t^{n,x})].$$

By using Remark 5.13 we see that $v_n(t,x) \to v(t,x)$ for each t and x.

By [90, Page 109] we have that

$$D_s X_t^{n,x} \frac{\partial}{\partial x} X_s^{n,x} = \frac{\partial}{\partial x} X_t^{n,x},$$

where the above product is the usual matrix product. So it follows that

$$\frac{\partial}{\partial x} X_t^{n,x} = \int_0^t a(s) D_s X_t^{n,x} \frac{\partial}{\partial x} X_s^{n,x} ds. \tag{5.41}$$

Interchanging integration and differentiation in connection with the chain rule we find that

$$\begin{aligned} \frac{\partial}{\partial x} v_n(t,x) &= E[\Phi'(X_t^{n,x}) \frac{\partial}{\partial x} X_t^{n,x}] \\ &= E[\int_0^t a(s) D_s \Phi(X_t^{n,x}) \frac{\partial}{\partial x} X_s^{n,x} ds] \\ &= E[\Phi(X_t^{n,x}) \int_0^t a(s) \left(\frac{\partial}{\partial x} X_s^{n,x} \right)^* dB_s]^*, \end{aligned}$$

where we applied the chain rule and the duality formula for the Malliavin derivative to the last equality.

Choose $\varphi \in C_0^\infty(U)$. In what follows, we will prove that

$$\int_{\mathbb{R}^d} \frac{\partial}{\partial x}\varphi(x)v(t,x)dx = -\int_{\mathbb{R}^d} \varphi(x)E[\Phi(X_t^x)\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^x\right)^* dB_s]^* dx. \tag{5.42}$$

In fact, dominated convergence combined with Remark 5.13 gives

$$\begin{aligned}\int_{\mathbb{R}^d} \frac{\partial}{\partial x}\varphi(x)v(t,x)dx &= -\lim_{n\to\infty}\int_{\mathbb{R}^d} \varphi(x)E[\Phi(X_t^{n,x})\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^{n,x}\right)^* dB_s]^* dx \\ &= -\lim_{n\to\infty}\int_{\mathbb{R}^d} \varphi(x)E[(\Phi(X_t^{n,x})-\Phi(X_t^x))\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^{n,x}\right)^* dB_s]^* dx \\ &\quad -\lim_{n\to\infty}\int_{\mathbb{R}^d} \varphi(x)E[\Phi(X_t^x)\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^{n,x}\right)^* dB_s]^* dx \\ &= -\lim_{n\to\infty} i)_n - \lim_{n\to\infty} ii)_n.\end{aligned}$$

As for the first term we get

$$i)_n \leq \int_{\mathbb{R}^d} |\varphi(x)|\|\frac{\partial}{\partial x}\Phi\|_\infty \|X_t^{n,x}-X_t^x\|_{L^2(\Omega;\mathbb{R}^d)}\|a\|_\infty \left(\sup_{k\geq 1, s\in[0,T]} E[\|\frac{\partial}{\partial x}X_s^{k,x}\|^2_{\mathbb{R}^{d\times d}}]\right)^{1/2} dx,$$

which goes to zero as n tends to infinity by Lebesque dominated convergence theorem, Remark 5.13 and Remark 5.8.

For the second term, $ii)_n$ since X_t^x is *Malliavin differentiable* and $\Phi \in C_b^2(\mathbb{R}^d)$ it follows from the Clark-Ocone formula that (see e.g. [90])

$$\Phi(X_t^x) = E[\Phi(X_t^x)] + \int_0^t E[D_s\Phi(X_t^x)|\mathcal{F}_s]dB_s.$$

So

$$\begin{aligned} ii)_n &= \int_{\mathbb{R}^d} \varphi(x)E[\Phi(X_t^x)\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^{n,x}\right)^* dB_s]^* dx \qquad (5.43)\\ &= \int_{\mathbb{R}^d} \varphi(x)E[\left(E[\Phi(X_t^x)] + \int_0^t E[D_s\Phi(X_t^x)|\mathcal{F}_s]dB_s\right)\int_0^t a(s)\left(\frac{\partial}{\partial x}X_s^{n,x}\right)^* dB_s]^* dx \\ &\qquad (5.44)\\ &= \int_0^t a(s)\int_{\mathbb{R}^d}\varphi(x)E[D_s\Phi(X_t^x)\frac{\partial}{\partial x}X_s^{n,x}]dxds. \qquad (5.45)\end{aligned}$$

One checks by means of Lemma 5.6 that $\varphi(\cdot)D_s\Phi(X_t^\cdot) = \varphi(\cdot)\Phi'(X_t^\cdot)D_sX_t^\cdot$ belongs to $L^2(\mathbb{R}^d \times \Omega;\mathbb{R}^d)$ so that for each s, the function

$$g_n(s) = \int_{\mathbb{R}^d}\varphi(x)E[D_s\Phi(X_t^x)\frac{\partial}{\partial x}X_s^{n,x}]dx$$

converges to $\int_{\mathbb{R}^d} \varphi(x) E[D_s \Phi(X_t^x) \frac{\partial}{\partial x} X_s^x] dx$ by the weak convergence of $\frac{\partial}{\partial x} X_s^{n,x}$ in $L^2([0,T] \times U \times \Omega)$ for a subsequence in virtue of Remark 5.8. Further,

$$\begin{aligned} |g_n(s)| &\leq \int_{\mathbb{R}^d} |\varphi(x)| \|D_s \Phi(X_t^x)\|_{L^2(\Omega;\mathbb{R}^d)} \|\frac{\partial}{\partial x} X_s^{n,x}\|_{L^2(\Omega;\mathbb{R}^d)} dx \\ &\leq \sup_{y \in \mathbb{R}^d,\, u \leq t,\, k \in \mathbb{N}} \|D_u \Phi(X_t^y)\|_{L^2(\Omega;\mathbb{R}^d)} \|\frac{\partial}{\partial x} X_u^{k,y}\|_{L^2(\Omega;\mathbb{R}^d)} \int_{\mathbb{R}^d} |\varphi(x)| dx \end{aligned}$$

so that Lebesgue's dominated convergence theorem gives

$$\lim_{n \to \infty} ii)_n = \int_0^t a(s) \int_{\mathbb{R}^d} \varphi(x) E[D_s \Phi(X_t^x) \frac{\partial}{\partial x} X_s^x] dx ds.$$

By reversing equations (5.43), (5.44) and (5.45) with $\frac{\partial}{\partial x} X_s^x$ in place of $\frac{\partial}{\partial x} X_s^{n,x}$ we obtain the result. $\square$

Appendix

5.A Framework

In this appendix we collect some facts from Gaussian white noise analysis and Malliavin calculus, which we shall use in Section 5.2 to construct strong solutions of SDE's. See [58, 94, 71] for more information on white noise theory. As for Malliavin calculus the reader may consult [90, 78, 79, 36].

5.A.1 Basic facts of Gaussian white noise theory

A crucial step in our proof for the construction of strong solutions (see Section 3) relies on a generalised stochastic process in the Hida distribution space which is shown to be a SDE solution. Let us first recall the definition of this space which is due to T. Hida (see [58]).

From now on we fix a time horizon $0 < T < \infty$. Let A be a (positive) self-adjoint operator on $L^2([0,T])$ with $Spec(A) > 1$. Require that A^{-r} is of Hilbert-Schmidt type for some $r > 0$ and let $\{e_j\}_{j\geq 0}$ be a complete orthonormal basis of $L^2([0,T])$ in $Dom(A)$ and let $\lambda_j > 0$, $j \geq 0$ be the eigenvalues of A such that

$$1 < \lambda_0 \leq \lambda_1 \leq \ldots \longrightarrow \infty.$$

Suppose that each basis element e_j is a continuous function on $[0,T]$. Further let O_λ, $\lambda \in \Gamma$, be an open covering of $[0,T]$ such that

$$\sup_{j\geq 0} \lambda_j^{-\alpha(\lambda)} \sup_{t\in O_\lambda} |e_j(t)| < \infty$$

for $\alpha(\lambda) \geq 0$.

In the sequel let $\mathcal{S}([0,T])$ be the standard countably Hilbertian space constructed from $(L^2([0,T]), A)$. See [94]. Then $\mathcal{S}([0,T])$ is a nuclear subspace of $L^2([0,T])$. The topological dual of $\mathcal{S}([0,T])$ is denoted by $\mathcal{S}'([0,T])$. Then the Bochner-Minlos theorem entails the existence of a unique probability measure π on $\mathcal{B}(\mathcal{S}'([0,T]))$ (Borel $\sigma-$algebra of $\mathcal{S}'([0,T])$) such that

$$\int_{\mathcal{S}'([0,T])} e^{i\langle \omega,\varphi\rangle} \pi(d\omega) = e^{-\frac{1}{2}\|\varphi\|^2_{L^2([0,T])}}$$

for all $\varphi \in \mathcal{S}([0,T])$, where $\langle \omega, \varphi \rangle$ stands for the action of $\omega \in \mathcal{S}'([0,T])$ on $\varphi \in \mathcal{S}([0,T])$. Define

$$\Omega_i = \mathcal{S}'([0,T])\,, \quad \mathcal{F}_i = \mathcal{B}(\mathcal{S}'([0,T]))\,, \quad \mu_i = \pi\,,$$

for $i = 1, \ldots, d$. Then the product measure

$$\mu = \mathop{\times}_{i=1}^{d} \mu_i \tag{5.46}$$

on the measurable space

$$(\Omega, \mathcal{F}) := \left(\prod_{i=1}^{d} \Omega_i, \bigotimes_{i=1}^{d} \mathcal{F}_i \right) \tag{5.47}$$

is called *d-dimensional white noise probability measure*.

Consider the Doléans-Dade exponential

$$\widetilde{e}(\varphi, \omega) = \exp\left(\langle \omega, \varphi \rangle - \frac{1}{2} \|\varphi\|^2_{L^2([0,T];\mathbb{R}^d)} \right),$$

for $\omega = (\omega_1, \ldots, \omega_d) \in (\mathcal{S}'([0,T]))^d$ and $\varphi = (\varphi^{(1)}, \ldots, \varphi^{(d)}) \in (\mathcal{S}([0,T]))^d$, where $\langle \omega, \varphi \rangle := \sum_{i=1}^{d} \langle \omega_i, \varphi_i \rangle$.

Now let $\left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} n}$ be the $n-$th completed symmetric tensor product of $(\mathcal{S}([0,T]))^d$ with itself. One checks that $\widetilde{e}(\varphi, \omega)$ is holomorphic in φ around zero. Hence, there exist generalised Hermite polynomials $H_n(\omega) \in \left(\left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} n}\right)'$ such that

$$\widetilde{e}(\varphi, \omega) = \sum_{n \geq 0} \frac{1}{n!} \left\langle H_n(\omega), \varphi^{\otimes n} \right\rangle \tag{5.48}$$

for φ in a certain neighbourhood of zero in $(\mathcal{S}([0,T]))^d$. One proves that

$$\left\{ \left\langle H_n(\omega), \varphi^{(n)} \right\rangle : \varphi^{(n)} \in \left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} n},\ n \in \mathbb{N}_0 \right\} \tag{5.49}$$

is a total set of $L^2(\Omega)$. Further, it can be shown that the generalised Hermite polynomials satisfy the orthogonality relation

$$\int_{\mathcal{S}'} \left\langle H_n(\omega), \varphi^{(n)} \right\rangle \left\langle H_m(\omega), \psi^{(m)} \right\rangle \mu(d\omega) = \delta_{n,m} n! \left(\varphi^{(n)}, \psi^{(n)} \right)_{L^2([0,T]^n;(\mathbb{R}^d)^{\otimes n})} \tag{5.50}$$

for all $n, m \in \mathbb{N}_0$, $\varphi^{(n)} \in \left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} n}$, $\psi^{(m)} \in \left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} m}$ where

$$\delta_{n,m} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{else} \end{cases}.$$

Denote by $\widehat{L}^2([0,T]^n; (\mathbb{R}^d)^{\otimes n})$ the space of square integrable symmetric functions $f(x_1, \ldots, x_n)$ with values in $(\mathbb{R}^d)^{\otimes n}$. Then it follows from relation (5.50) that the mappings

$$\varphi^{(n)} \longmapsto \left\langle H_n(\omega), \varphi^{(n)} \right\rangle$$

from $\left(\mathcal{S}([0,T])^d\right)^{\widehat{\otimes} n}$ to $L^2(\Omega)$ have unique continuous extensions

$$I_n : \widehat{L}^2([0,T]^n;(\mathbb{R}^d)^{\otimes n}) \longrightarrow L^2(\Omega)$$

for all $n \in \mathbb{N}$. These extensions $I_n(\varphi^{(n)})$ can be identified as n-fold iterated Itô integrals of $\varphi^{(n)} \in \widehat{L}^2([0,T]^n;(\mathbb{R}^d)^{\otimes n})$ with respect to a $d-$dimensional Wiener process

$$B_t = \left(B_t^{(1)}, \ldots, B_t^{(d)}\right) \tag{5.51}$$

on the white noise space

$$(\Omega, \mathcal{F}, \mu) \ . \tag{5.52}$$

We mention that square integrable functionals of B_t admit a Wiener-Itô chaos representation which can be regarded as an infinite-dimensional Taylor expansion, that is

$$L^2(\Omega) = \bigoplus_{n \geq 0} I_n(\widehat{L}^2([0,T]^n;(\mathbb{R}^d)^{\otimes n})). \tag{5.53}$$

The definition of the Hida stochastic test function and distribution space is based on the Wiener-Itô chaos decomposition (5.53): Set

$$A^d := (A, \ldots, A)\,. \tag{5.54}$$

Using a second quantisation argument, the Hida stochastic test function space $(\mathcal{S})$ is defined as the space of all $f = \sum_{n \geq 0} \left\langle H_n(\cdot), \varphi^{(n)} \right\rangle \in L^2(\Omega)$ such that

$$\|f\|_{0,p}^2 := \sum_{n \geq 0} n! \left\| \left((A^d)^{\otimes n}\right)^p \varphi^{(n)} \right\|_{L^2([0,T]^n;(\mathbb{R}^d)^{\otimes n})}^2 < \infty \tag{5.55}$$

for all $p \geq 0$. In fact, the space $(\mathcal{S})$ is a nuclear Fréchet algebra with respect to multiplication of functions and its topology is induced by the seminorms $\|\cdot\|_{0,p}$, $p \geq 0$. Further one shows that

$$\widetilde{e}(\varphi, \omega) \in (\mathcal{S}) \tag{5.56}$$

for all $\varphi \in (\mathcal{S}([0,T]))^d$.

On the other hand, the topological dual of $(\mathcal{S})$, denoted by $(\mathcal{S})^*$, is called *Hida stochastic distribution space*. Using these definitions we ontain the Gel'fand triple

$$(\mathcal{S}) \hookrightarrow L^2(\Omega) \hookrightarrow (\mathcal{S})^*.$$

It turns out that the *white noise* of the coordinates of the $d-$dimensional Wiener process B_t, that is the time derivatives

$$W_t^i := \frac{d}{dt} B_t^i,\ i = 1, \ldots, d\,, \tag{5.57}$$

belong to $(\mathcal{S})^*$.

We also recall the definition of the S-transform. See [98]. The $S-$transform of a $\Phi \in (\mathcal{S})^*$, denoted by $S(\Phi)$, is defined by the dual pairing

$$S(\Phi)(\varphi) = \langle \Phi, \widetilde{e}(\varphi, \omega) \rangle \tag{5.58}$$

for $\varphi \in (\mathcal{S}_{\mathbb{C}}([0,T]))^d$. Here $\mathcal{S}_{\mathbb{C}}([0,T])$ the complexification of $\mathcal{S}([0,T])$. The $S-$transform is a monomorphism from $(\mathcal{S})^*$ to $\mathbb{C}$. In particular, if

$$S(\Phi) = S(\Psi) \text{ for } \Phi, \Psi \in (\mathcal{S})^*$$

then

$$\Phi = \Psi.$$

As an example one finds that

$$S(W_t^i)(\varphi) = \varphi^i(t), \ i = 1, ..., d \tag{5.59}$$

for $\varphi = (\varphi^{(1)}, \dots, \varphi^{(d)}) \in (\mathcal{S}_{\mathbb{C}}([0,T]))^d$.

Finally, we recall the concept of the *Wick* or *Wick-Grassmann product*. The Wick product defines a tensor algebra multiplication on the Fock space and is introduced as follows: The Wick product of two distributions $\Phi, \Psi \in (\mathcal{S})^*$, denoted by $\Phi \diamond \Psi$, is the unique element in $(\mathcal{S})^*$ such that

$$S(\Phi \diamond \Psi)(\varphi) = S(\Phi)(\varphi) S(\Psi)(\varphi) \tag{5.60}$$

for all $\varphi \in (\mathcal{S}_{\mathbb{C}}([0,T]))^d$. As an example, we get

$$\left\langle H_n(\omega), \varphi^{(n)} \right\rangle \diamond \left\langle H_m(\omega), \psi^{(m)} \right\rangle = \left\langle H_{n+m}(\omega), \varphi^{(n)} \widehat{\otimes} \psi^{(m)} \right\rangle \tag{5.61}$$

for $\varphi^{(n)} \in \left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} n}$ and $\psi^{(m)} \in \left((\mathcal{S}([0,T]))^d\right)^{\widehat{\otimes} m}$. The latter in connection with (5.48) implies that

$$\widetilde{e}(\varphi, \omega) = \exp^{\diamond}(\langle \omega, \varphi \rangle) \tag{5.62}$$

for $\varphi \in (\mathcal{S}([0,T]))^d$. Here the Wick exponential $\exp^{\diamond}(X)$ of a $X \in (\mathcal{S})^*$ is defined as

$$\exp^{\diamond}(X) = \sum_{n \geq 0} \frac{1}{n!} X^{\diamond n}, \tag{5.63}$$

where $X^{\diamond n} = X \diamond \ldots \diamond X$, provided that the sum on the right hand side converges in $(\mathcal{S})^*$.

5.A.2 Basic elements of Malliavin Calculus

In this section we pass in review some basic definitions from Malliavin calculus.

For convenience we consider the case $d = 1$. Let $F \in L^2(\Omega)$. Then we know from (5.53) that

$$F = \sum_{n \geq 0} \left\langle H_n(\cdot), \varphi^{(n)} \right\rangle \tag{5.64}$$

for unique $\varphi^{(n)} \in \widehat{L}^2([0,T]^n)$. Suppose that

$$\sum_{n\geq 1} nn! \left\|\varphi^{(n)}\right\|^2_{L^2([0,T]^n)} < \infty\,. \tag{5.65}$$

Then the *Malliavin derivative* D_t of F in the direction of B_t can be defined as

$$D_t F = \sum_{n\geq 1} n \left\langle H_{n-1}(\cdot), \varphi^{(n)}(\cdot, t)\right\rangle. \tag{5.66}$$

We denote by $\mathbb{D}^{1,2}$ the space of all $F \in L^2(\Omega)$ such that (5.65) holds. The Malliavin derivative $D.$ is a linear operator from $\mathbb{D}^{1,2}$ to $L^2([0,T]\times\Omega)$. We mention that $\mathbb{D}^{1,2}$ is a Hilbert space with the norm $\|\cdot\|_{1,2}$ given by

$$\|F\|^2_{1,2} := \|F\|^2_{L^2(\Omega,\mu)} + \|D.F\|^2_{L^2([0,T]\times\Omega,\lambda\times\mu)}\,. \tag{5.67}$$

We get the following chain of continuous inclusions:

$$(\mathcal{S}) \hookrightarrow \mathbb{D}^{1,2} \hookrightarrow L^2(\Omega) \hookrightarrow \mathbb{D}^{-1,2} \hookrightarrow (\mathcal{S})^*, \tag{5.68}$$

where $\mathbb{D}^{-1,2}$ is the dual of $\mathbb{D}^{1,2}$.

5.B Technical results

We give a list if technical results needed for the proofs of Section 5.2 and 5.3.

Lemma 5.17. *Let $\{f_n\}_{n\geq 0}$ be a bounded sequence of functions in L^q_p. Then, for every $k \in \mathbb{R}$*

$$\sup_{x\in\mathbb{R}^d}\sup_{n\geq 0} E\left[\exp\left\{k\int_0^T |f_n(s,x+B_s)|^2 ds\right\}\right] < \infty.$$

In particular, there exists a weak solution to SDE (5.5).

Proof. See [68, Lemma 3.2] □

Lemma 5.18. *Let $\{f_n\}_{n\geq 0}$ a sequence of elements in L^q_p that converges to some $f \in L^q_p$. Then there exists $\varepsilon > 1$ such that*

$$\sup_{n\geq 0} E\left[\int_0^T \|f_n(s,\varphi^n_s)\|^{2\varepsilon} ds\right] < \infty. \tag{5.69}$$

Here $\varphi^n_s : x \mapsto X^{x,n}_t$ denotes the stochastic flow associated to the solution of the SDE (5.5) with drift coefficient $b_n \in C^\infty_b(\mathbb{R}^d)$.

Proof. See [44, Lemma 15]. □

We also need the following crucial lemma, which can be found in [44], Lemma 3.4.

Lemma 5.19. *Let U_n be the solution of the PDE (5.10) with $\Phi = b = b_n \in C^\infty_b(\mathbb{R}^n)$. Let $X^{x,n}_t$ be the solution of the SDE (5.5) with drift coefficient $b_n \in C^\infty_b(\mathbb{R}^d)$. Then the following holds true* *(i) For each* $r > 0$ *there exists a function* f *with* $\lim_n f(n) = 0$ *such that*

$$\sup_{x \in B_r} \sup_{t \in [0,T]} \|U_n(t,x) - U(t,x)\| \leq f(n)$$

and

$$\sup_{x \in B_r} \sup_{t \in [0,T]} \|\nabla U_n(t,x) - \nabla U(t,x)\| \leq f(n)$$

(ii) There exists a $\lambda \in \mathbb{R}$ *for which* $\sup_{\substack{t \in [0,T] \\ x \in \mathbb{R}^d}} \|\nabla U_n(t,x)\| \leq \frac{1}{2}$.

(iii) $\sup_{n \geq 0} \|\Delta U_n(t,x)\|_{L_p^q} < \infty$.

(iv) As a consequence of the boundedness of U_n *and* ∇U_n *we have*

$$\sup_{t \in [0,T]} E\left[\|\gamma_t^n(x)\|^a\right] \leq C\left(1 + |x|^a\right).$$

The following lemma gives a bound for the derivative of the inverse of the family of diffeomorphisms γ_t. See [44], Lemma 3.5 for its proof.

Lemma 5.20. *Let* $\gamma_{t,n} : \mathbb{R}^d \to \mathbb{R}^d$ *be the* C^1*-diffeomorphisms defined as* $\gamma_{t,n}(x) := x + U_n(t,x)$ *for* $x \in \mathbb{R}^d$ *associated to* $X_t^{x,n}$ *the solution of SDE (5.5) with drift coefficient* $b_n \in C_b^\infty(\mathbb{R}^d)$. *Then*

$$\sup_{n \geq 0} \sup_{t \in [0,T]} \|\nabla \gamma_{t,n}^{-1}\|_{C(\mathbb{R}^d)} \leq 2.$$

The next result was shown in [43], Corollary 13.

Lemma 5.21. *Let* V_t^n *be the process defined in (5.22). Then for every* $\alpha \in \mathbb{R}$

$$\sup_{n \geq 0} E\left[e^{\alpha V_T^n}\right] \leq C.$$

Observe that the same estimate holds for any $t \in [0,T]$ *since* V_t^n *is an increasing process.*

5.C A compactness criterion for subsets of $L^2(\Omega)$

The following result which is due to [30, Theorem 1] gives a compactness criterion for subsets of $L^2(\Omega; \mathbb{R}^d)$ using Malliavin calculus.

Theorem 5.22. *Let* $\{(\Omega, \mathcal{A}, P); H\}$ *be a Gaussian probability space, that is* $(\Omega, \mathcal{A}, P)$ *is a probability space and* H *a separable closed subspace of Gaussian random variables of* $L^2(\Omega)$, *which generate the* σ*-field* $\mathcal{A}$. *Denote by* $\mathbf{D}$ *the derivative operator acting on elementary smooth random variables in the sense that*

$$\mathbf{D}(f(h_1, \ldots, h_n)) = \sum_{i=1}^n \partial_i f(h_1, \ldots, h_n) h_i, \ h_i \in H, f \in C_b^\infty(\mathbb{R}^n).$$

Further let $\mathbf{D}_{1,2}$ be the closure of the family of elementary smooth random variables with respect to the norm

$$\|F\|_{1,2} := \|F\|_{L^2(\Omega)} + \|\mathbf{D}F\|_{L^2(\Omega;H)} .$$

Assume that C is a self-adjoint compact operator on H with dense image. Then for any $c > 0$ the set

$$\mathcal{G} = \left\{ G \in \mathbf{D}_{1,2} : \|G\|_{L^2(\Omega)} + \left\|C^{-1}\mathbf{D}\, G\right\|_{L^2(\Omega;H)} \leq c \right\}$$

is relatively compact in $L^2(\Omega)$.

A useful bound in connection with Theorem 5.22, based on fractional Sobolev spaces is the following (see [30]):

Lemma 5.23. *Let $v_s, s \geq 0$ be the Haar basis of $L^2([0,T])$. For any $0 < \alpha < 1/2$ define the operator A_α on $L^2([0,T])$ by*

$$A_\alpha v_s = 2^{k\alpha} v_s, \text{ if } s = 2^k + j$$

for $k \geq 0, 0 \leq j \leq 2^k$ and

$$A_\alpha T = T.$$

Then for all β with $\alpha < \beta < (1/2)$, there exists a constant c_1 such that

$$\|A_\alpha f\| \leq c_1 \left\{ \|f\|_{L^2([0,T])} + \left(\int_0^T \int_0^T \frac{|f(t) - f(t')|^2}{|t - t'|^{1+2\beta}} dt\, dt' \right)^{1/2} \right\}.$$

A direct consequence of Theorem 5.22 and Lemma 5.23 is now the following compactness criterion which is essential for the proof of Corollary 5.9.

Corollary 5.24. *Let a sequence of $\mathcal{F}_T$-measurable random variables $X_n \in \mathbb{D}_{1,2}$, $n = 1, 2...$, be such that there exist constants $\alpha > 0$ and $C > 0$ with*

$$\sup_n E[|X_n|^2] \leq C,$$

$$\sup_n E\left[\|D_t X_n - D_{t'} X_n\|^2\right] \leq C|t - t'|^\alpha$$

for $0 \leq t' \leq t \leq T$ and

$$\sup_n \sup_{0 \leq t \leq T} E\left[\|D_t X_n\|^2\right] \leq C .$$

Then the sequence X_n, $n = 1, 2...$, is relatively compact in $L^2(\Omega)$.

Chapter 6

Computing Deltas without derivatives

David R. Baños, Sindre Duedahl, Thilo Meyer-Brandis and Frank Proske

Abstract: A well-known application of Malliavin calculus in Mathematical Finance is the probabilistic representation of option price sensitivities, the so-called Greeks, as expectation functionals that do not involve the derivative of the pay-off function. This allows for numerically tractable computation of the Greeks even for discontinuous pay-off functions. However, while the pay-off function is allowed to be irregular, the coefficients of the underlying diffusion are required to be smooth in the existing literature, which for example excludes already simple regime switching diffusion models. The aim of this article is to generalise this application of Malliavin calculus to Itô diffusions with irregular drift coefficients, whereat we here focus on the computation of the Delta, which is the option price sensitivity with respect to the initial value of the underlying. To this purpose we first show existence, Malliavin differentiability, and (Sobolev) differentiability in the initial condition of strong solutions of Itô diffusions with drift coefficients that can be decomposed into the sum of a bounded but merely measurable and a Lipschitz part. Furthermore, we give explicit expressions for the corresponding Malliavin and Sobolev derivative in terms of the local time of the diffusion, respectively. We then turn to the main objective of this article and analyse the existence and probabilistic representation of the corresponding Deltas for lookback and Asian type options. We conclude with a simulation study of several regime-switching examples.

6.1 Introduction

Throughout this paper, let $T > 0$ be a given time horizon and $(\Omega, \mathcal{F}, P)$ a complete probability space equipped with a one-dimensional Brownian motion $\{B_t\}_{t\in[0,T]}$ and the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ generated by $\{B_t\}_{t\in[0,T]}$ augmented by the P-null sets. Further, we will only deal with random variables that are Brownian functionals, i.e. we assume $\mathcal{F} := \mathcal{F}_T$.

One of the most prominent applications of Malliavin calculus in financial mathematics concerns the derivation of numerically tractable expressions for the so-called *Greeks*, which are important sensitivities of option prices with respect to involved parameters. The first paper to address this application was [50], which has consecutively triggered an active research interest in this topic, see e.g. [49], [19], [4]. See also [27], [38] and references therein for a related approach based on functional Itô calculus. Suppose the risk-neutral dynamics of the underlying asset of a European option is driven by a stochastic differential equation (for short SDE) of the form

$$dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dB_t, \;\; X_0^x = x \in \mathbb{R}\,,$$

where $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ are some given drift and volatility coefficients, respectively. Let $\Phi : \mathbb{R} \to \mathbb{R}$ denote the pay-off function and the expectation $E[\Phi(X_T^x)]$ the risk-neutral price at time zero of the option with maturity $T > 0$. For notational simplicity we assume the discounting rate to be zero. In this paper we will focus on the *Delta*

$$\frac{\partial}{\partial x} E[\Phi(X_T^x)]\,, \tag{6.1}$$

which is a measure for the sensitivity of the option price with respect to changes of the initial value of the underlying asset. As is well known, the Delta has a particular role among the Greeks as it determines the hedge portfolio in many complete market models. If the drift $b(\cdot)$, the volatility $\sigma(\cdot)$, and the pay-off $\Phi(\cdot)$ are "sufficiently regular" to allow for differentiation under the expectation, the Delta can be computed in a straight-forward manner as

$$E\left[\frac{\partial}{\partial x}\Phi(X_T^x)\right] = E[\Phi'(X_T^x)\, Z_T]\,, \tag{6.2}$$

where the *first variation process* $Z_t := \frac{\partial}{\partial x} X_t^x$ is given by

$$Z_t = \exp\left\{\int_0^t \left[b'(X_s^x) - \frac{1}{2}(\sigma'(X_s^x))^2\right] ds + \int_0^t \sigma'(X_s^x)\, dB_s\right\}\,, \tag{6.3}$$

and where Φ', b', σ' denote the derivatives of Φ, b, σ, respectively. For example, requiring that Φ, b, σ are continuously differentiable with bounded derivatives would allow (6.2) to hold (we refer to [70] for conditions on b and σ that guarantee the existence of the first variation process), and the expectation in (6.2) could be approximated e.g. by Monte Carlo methods. In most realistic situations, though, straight-forward computations as in (6.2) are not possible. In that case, one could combine numerical methods to approximate the derivative and the expectation in (6.1), respectively, to compute the Delta. However, in particular for discontinuous pay-offs Φ as is the case for a digital option this procedure might be numerically inefficient, see for example [50]. At that point, the following result for lookback options obtained with the help of Malliavin calculus appears to be useful, where the option pay-off is allowed to depend on the path of the underlying at finitely many time points.

Theorem 6.1 (Proposition 3.2 in [50]). *Let $b(\cdot)$ and $\sigma(\cdot)$ be continuously differentiable with bounded Lipschitz derivatives, $\sigma(\cdot) > \epsilon > 0$, and $\Phi : \mathbb{R}^m \to \mathbb{R}$ be such that the pay-off $\Phi(X_{T_1}^x, \ldots, X_{T_m}^x)$, $T_1, \ldots, T_m \in (0, T]$, of the corresponding lookback option is square integrable. Then the Delta exists and is given by*

$$\frac{\partial}{\partial x} E[\Phi(X_{T_1}^x, \ldots, X_{T_m}^x)] = E\left[\Phi(X_{T_1}^x, \ldots, X_{T_m}^x) \int_0^T a(t)\, \sigma^{-1}(X_t^x)\, Z_t\, dB_t\right]\,, \tag{6.4}$$

where Z_t is the first variation process given in (6.3) *and $a(t)$ is any square integrable determin-*

istic function such that, for every $i = 1, \dots, m$,

$$\int_0^{T_i} a(s)ds = 1.$$

While for notational simplicity we present the above result for one-dimensional X^x we remark that in [50] the extension to multi-dimensional underlying asset and Brownian motion is considered. If the option is of European type, i.e. the pay-off $\Phi(X_T^x)$ depends only on the underlying at T, then (6.4) is the probabilistic representation of the space derivative of a solution to a Kolmogorov equation which is also referred to as *Bismuth-Elworthy-Li type formula* in the literature due to [40], [21]. The strength of (6.4) is that the Delta is expressed again as an expectation of the pay-off multiplied by the so-called Malliavin weight $\int_0^T a(t)\,\sigma^{-1}(X_t^x)\,Z_t\,dB_t$. Computing the Delta by Monte-Carlo via this reformulation then guarantees a convergence rate that is independent of the regularity of the pay-off function Φ and the dimensionality. Note that the Malliavin weight is independent of the option pay-off, and thus the same weight can be employed in the computations of the Deltas of different options. Also, in [49] and [18] the question of how to optimally choose the function $a(t)$ with respect to computational efficiency is considered.

While the representation (6.4) succeeds to handle irregular pay-offs by getting rid of the derivative of Φ, the regularity assumptions on the coefficients b and σ driving the dynamics of the underlying diffusion are rather strong. Consider for example an extended Black and Scholes model where the stock pays a dividend yield that switches to a higher level when the stock value passes a certain threshold. Then, again with the risk-free rate equal to zero for simplicity, the logarithm of the stock price is modelled by the following dynamics under the risk-neutral measure:

$$dX_t^x = b(X_t^x)dt + \sigma dB_t, \;\; X_0^x = x \in \mathbb{R}\,,$$

where $\sigma > 0$ is constant and the drift coefficient $b : \mathbb{R} \to \mathbb{R}$ is given by

$$b(x) := -\lambda_1 \mathbf{1}_{(-\infty,R)}(x) - \lambda_2 \mathbf{1}_{[R,\infty)}(x),$$

for dividend yields $\lambda_1, \lambda_2 \in \mathbb{R}_+$ and a given threshold $R \in \mathbb{R}$. In [35], a (more complex) irregular drift b is interpreted as state-dependent fees deducted by the insurer in the evolution of variable annuities instead of dividend yield. Already, this simple regime-switching model is not covered by the result in Theorem 6.1 since the drift coefficient is not continuously differentiable.

Or allow for state-dependent regime-switching of the mean reversion rate in an extended Ornstein-Uhlenbeck process:

$$dX_t^x = b(X_t^x)dt + \sigma dB_t, \;\; X_0^x = x \in \mathbb{R}\,,$$

where $\sigma > 0$ is constant and the drift coefficient $b : \mathbb{R} \to \mathbb{R}$ is given by

$$b(x) := -\lambda_1 x \mathbf{1}_{(-\infty,R)}(x) - \lambda_2 x \mathbf{1}_{[R,\infty)}(x)$$

for mean reversion rates $\lambda_1, \lambda_2 \in \mathbb{R}_+$ and a given threshold $R \in \mathbb{R}$ (here the mean reversion level is set equal to zero). This type of model captures well, for instance, the evolution of electricity spot prices, which switches between so-called spike regimes on high price levels with very fast mean reversion and base regimes on normal price levels with moderate speed of mean reversion, see e.g. [20], [63], [85] and references therein. Alternatively, an extended Ornstein-Uhlenbeck process with state-dependent regime-switching of the mean reversion level (low and high interest rate environments) is an interesting modification of the Vašíček short rate model. Note that in that case the Delta is rather a generalised Rho, i.e. a sensitivity measure with respect to the short end of the yield curve. We observe that also these two extended Ornstein-Uhlenbeck processes are not covered by the result in Theorem 6.1.

Motivated by these examples, this paper aims at deriving an analogous result to Theorem 6.1 when the underlying is driven by an SDE with irregular drift coefficient. More precisely, we will consider SDE's

$$dX_t^x = b(t, X_t^x)dt + dB_t,\ 0 \le t \le T,\ X_0^x = x \in \mathbb{R}\,, \tag{6.5}$$

where we allow for time-inhomogeneous drift coefficients $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ in the form

$$b(t, x) = \tilde{b}(t, x) + \hat{b}(t, x)\,,\ (t, x) \in [0, T] \times \mathbb{R}\,, \tag{6.6}$$

for $\tilde{b}$ merely bounded and measurable, and $\hat{b}$ Lipschitz continuous and at most of linear growth in x uniformly in t, i.e. there exists a constant $C > 0$ such that

$$|\hat{b}(t, x) - \hat{b}(t, y)| \le C|x - y| \tag{6.7}$$

$$|\hat{b}(t, x)| \le C(1 + |x|) \tag{6.8}$$

for $x, y \in \mathbb{R}$ and $t \in [0, T]$. Adding the Lipschitz component $\hat{b}(t, x)$ in (6.6) is motivated by the fact that many drift coefficients interesting for financial applications are of linear growths. At present we are not able to show our results for general measurable drift coefficients of linear growths, but only for those where the irregular behavior remains in a bounded spectrum. However, from an application point of view this class is very rhich already, and in particular it contains the regime switching examples from above. In (6.5) we consider a constant volatility coefficient $\sigma(t, x) := 1$, but we will see at the end of Section 6.3 (Theorem 6.21) that our results apply to many SDE's with more general volatility coefficients which can be reduced to SDE's of type (6.5) (which for example is possible for volatility coefficients as in Theorem 6.1).

In order to be able to apply Malliavin calculus to the underlying diffusion, the first thing we need to ensure is that the solution of SDE (6.5) is a Brownian functionals, i.e. we are interested in the existence of strong solutions of (6.5).

Definition 6.2. *A* strong solution *of SDE* (6.5) *is a continuous,* $\{\mathcal{F}_t\}_{t\in[0,T]}$*-adapted process* $\{X_t^x\}_{t\in[0,T]}$ *that solves equation* (6.5)*.*

Remark 6.3. *Note that the usual definition of a strong solution requires the existence of a Brownian-adapted solution of* (6.5) *on* any *given stochastic basis. However, an* $\{\mathcal{F}_t\}_{t\in[0,T]}$*-adapted solution* $\{X_t^x\}_{t\in[0,T]}$ *on the given stochastic basis* $(\Omega, \mathcal{F}, P, B)$ *can be written in the form* $X_t^x = F_t(B_\cdot)$ *for some family of functionals* F_t*,* $t \in [0, T]$*, (see e.g. [82] for an explicit form of* F_t*). Then for any other stochastic basis* $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \hat{B})$ *one gets that* $X_t^x := F_t(\hat{B}_\cdot)$*,* $t \in [0, T]$*, is a* $\hat{B}$*-adapted solution to SDE (6.5). So once there is a Brownian-adapted solution* *of* (6.5) *on one given stochastic basis, it follows that there indeed exists a strong solution in the usual sense. This justifies our definition of a strong solution above.*

To pursue our objectives we proceed as follows in the remaining parts of the paper. In Section 6.2 we recall some fundamental concepts from Malliavin calculus and local time calculus which compose central mathematical tools in the following analysis.

We then analyse in Section 6.3 the existence and Malliavin differentiability of a unique strong solution of SDE's with irregular drift coefficients as in (6.5) (Theorem 6.14). It is well known that the SDE is Malliavin differentiable as soon as the coefficients are Lipschitz continuous (see e.g. [90]); for merely bounded and measurable drift coefficients Malliavin differentiability was shown only recently in [83], (see also [81]). Here, we extend ideas introduced for bounded coefficients in [83] to drift coefficients of type (6.6). Unlike in most of the existing literature on strong solutions of SDE's with irregular coefficients our approach does not rely on a pathwise uniqueness argument (Yamada-Watanabe Theorem). Instead, we employ a compactness criterium based on Malliavin calculus together with local time calculus to directly construct a strong solution which in addition is Malliavin differentiable. Also, we are able to give an explicit expression for the Malliavin derivative of the strong solution of (6.5) in terms of the integral of b (and not the derivative of b) with respect to local time of the strong solution (Proposition 6.15). We mention that while existence and Malliavin differentiability of strong solutions could be extended to analogue multi-dimensional SDE's as in [81], the explicit expression of the Malliavin derivative is in general only possible for one-dimensional SDE's as considered in this paper. Moreover, in this paper we replace arguments that are based on White Noise analysis in [83] and [81] by alternative proofs which might make the text more accessible for readers who are unfamiliar with concepts from White Noise analysis.

Next, we need to analyse the regularity of the dependence of the strong solution in its initial condition and to introduce the analogue of the first variation process (6.3) in case of irregular drift coefficients. Using the close connection between the Malliavin derivative and the first variation process, we find that the strong solution is Sobolev differentiable in its initial condition (Theorem 6.17). Again, we give an explicit expression for the corresponding (Sobolev) first variation process which does not include the derivative of b (Proposition 6.18).

In Section 6.4 we develop our main result (Theorem 6.23) which extends Theorem 6.1 to SDE's with irregular drift coefficients. To this end, one has to show in the first place that the Delta exists, i.e. that $E[\Phi(X^x_{T_1}, \ldots, X^x_{T_m})]$ is continuously differentiable in x. At this point the explicit expressions for the Malliavin derivative and the first variation process are essential. In the final representation of the Delta we then have gotten rid of both the derivative of the pay-off Φ and the derivative of the drift coefficient b in the first variation process, whence the title "Computing Deltas without Derivatives" of the paper. In addition to Deltas of lookback options as in Theorem 6.1, we further consider Deltas of Asian options with pay-offs of the type $\Phi\left(\int_{T_1}^{T_2} X^x_u \, du\right)$ for $T_1, T_2 \in [0, T]$ and some function $\Phi : \mathbb{R} \to \mathbb{R}$. In case the starting point of the averaging period of the Asian pay-off lies in the future, i.e. $T_1 > 0$, we are able to give analogue results to the ones of lookback options. If the averaging period starts today, i.e. $T_1 = 0$, the Malliavin weight in the expression for the Delta would include a general Skorohod integral which is neither numerically nor mathematically tractable in our analysis (except for linear coefficients as in the Black and Scholes model where the Skorohod integral turns out to be an

Itô integral). However, we are still able to state two approximation results for the Delta in this case.

In Section 6.5 we consider some examples and compute the Deltas in the concrete regime-switching models mentioned above. We do a small simulation study and compare the performance to a finite difference approximation of the Delta in the same spirit as in [50].

We conclude the paper by an appendix with some technical proofs from Section 6.3 which have been deferred to the end of the paper for better readability.

Notations: We summarise some of the most frequently used notations:

- $C^1(\mathbb{R})$ denotes the space of continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$.
- $C_0^\infty([0,T] \times \mathbb{R})$, respectively $C_0^\infty(\mathbb{R})$, denotes the space of infinitely many times differentiable functions on $[0,T] \times \mathbb{R}$, respectively $\mathbb{R}$, with compact support.
- For a measurable space $(S, \mathcal{G})$ equipped with a measure μ, we denote by $L^p(S, \mathcal{G})$ or $L^p(S)$ the Banach space of (equivalence classes of) functions on S integrable to some power p, $p \geq 1$.
- $L^p_{loc}(\mathbb{R})$ denotes the space of locally Lebesgue integrable functions to some power p, $p \geq 1$, i.e. $\int_U |f(x)|^p dx < \infty$ for every open bounded subset $U \subset \mathbb{R}$.
- $W^{1,p}_{loc}(\mathbb{R})$ denotes the subspace of $L^p_{loc}(\mathbb{R})$ of weakly (Sobolev) differentiable functions such that the weak derivative f' belongs to $L^p_{loc}(\mathbb{R})$, $p \geq 1$.
- For a progressive process Y. we denote the Doléans-Dade exponential of the corresponding Brownian integral (if well defined) by

$$\mathcal{E}\left(\int_0^t b(u, Y_u) dB_u\right) := \exp\left(\int_0^t b(u, Y_u) dB_u - \frac{1}{2}\int_0^t b^2(u, Y_u) du\right), \quad t \in [0,T]. \tag{6.9}$$

- For $Z \in L^2(\Omega, \mathcal{F}_T)$ we denote the *Wiener-transform* of Z in $f \in L^2([0,T])$ by

$$\mathcal{W}(Z)(f) := E\left[Z\mathcal{E}\left(\int_0^T f(s) dB_s\right)\right].$$

- We will use the symbol $\lesssim$ to denote *less or equal than* up to a positive real constant $C > 0$ not depending on the parameters of interest, i.e. if we have two mathematical expressions $E_1(\theta), E_2(\theta)$ depending on some parameter of interest θ then $E_1(\theta) \lesssim E_2(\theta)$ if, and only if, there is a positive real number $C > 0$ independent of θ such that $E_1(\theta) \leq C E_2(\theta)$.

6.2 Framework

Our main results centrally rely on tools from Malliavin calculus as well as integration with respect to local time both in time and space. We here provide a concise introduction to the main concepts in these two areas that will be employed in the following sections. For deeper information on Malliavin calculus the reader is referred to i.e. [90, 78, 79, 36]. As for theory on local time integration for Brownian motion we refer to i.e. [39, 100].

6.2.1 Malliavin calculus

Denote by $\mathcal{S}$ the set of simple random variables $F \in L^2(\Omega)$ in the form

$$F = f\left(\int_0^T h_1(s)dB_s, \ldots, \int_0^T h_n(s)dB_s\right), \quad h_1, \ldots, h_n \in L^2([0,T]), \;\; f \in C_0^\infty(\mathbb{R}^n).$$

The Malliavin derivative operator D acting on such simple random variables is the process $DF = \{D_tF, t \in [0,T]\}$ in $L^2(\Omega \times [0,T])$ defined by

$$D_tF = \sum_{i=1}^n \partial_i f\left(\int_0^T h_1(s)dB_s, \ldots, \int_0^T h_n(s)dB_s\right) h_i(t).$$

Define the following norm on $\mathcal{S}$:

$$\|F\|_{1,2} := \|F\|_{L^2(\Omega)} + \|DF\|_{L^2(\Omega;L^2([0,T]))} = E[|F|^2]^{1/2} + E\left[\int_0^T |D_tF|^2 dt\right]^{1/2}. \tag{6.10}$$

We denote by $\mathbb{D}^{1,2}$ the closure of the family of simple random variables $\mathcal{S}$ with respect to the norm given in (6.10), and we will refer to this space as the space of Malliavin differentiable random variables in $L^2(\Omega)$ with Malliavin derivative belonging to $L^2(\Omega)$.

In the derivation of the probabilistic representation for the Delta, the following chain rule for the Malliavin derivative will be essential:

Lemma 6.4. *Let $\varphi : \mathbb{R}^m \to \mathbb{R}$ be continuously differentiable with bounded partial derivatives. Further, suppose that $F = (F_1, \ldots, F_m)$ is a random vector whose components are in $\mathbb{D}^{1,2}$. Then $\varphi(F) \in \mathbb{D}^{1,2}$ and*

$$D_t\varphi(F) = \sum_{i=1}^m \partial_i\varphi(F)D_tF_i, \quad P-a.s., \quad t \in [0,T].$$

The Malliavin derivative operator $D : \mathbb{D}^{1,2} \to L^2(\Omega \times [0,T])$ admits an adjoint operator $\delta = D^* : \mathrm{Dom}(\delta) \to L^2(\Omega)$ where the domain $\mathrm{Dom}(\delta)$ is characterised by all $u \in L^2(\Omega \times [0,T])$ such that for all $F \in \mathbb{D}^{1,2}$ we have

$$E\left[\int_0^T D_tF\, u_t dt\right] \le C\|F\|_{1,2},$$

where C is some constant depending on u.

For a stochastic process $u \in \mathrm{Dom}(\delta)$ (not necessarily adapted to $\{\mathcal{F}_t\}_{t\in[0,T]}$) we denote by

$$\delta(u) := \int_0^T u_t \delta B_t \tag{6.11}$$

the action of δ on u. The above expression (6.11) is known as the Skorokhod integral of u and it is an anticipative stochastic integral. It turns out that all $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted processes in $L^2(\Omega \times [0,T])$ are in the domain of δ and for such processes u_t we have

$$\delta(u) = \int_0^T u_t dB_t,$$

i.e.the Skorokhod and Itô integrals coincide. In this sense, the Skorokhod integral can be considered to be an extension of the Itô integral to non-adapted integrands.

The dual relation between the Malliavin derivative and the Skorokhod integral implies the following important formula:

Theorem 6.5 (Duality formula)**.** *Let $F \in \mathbb{D}^{1,2}$ and $u \in$ Dom(δ). Then*

$$E\left[F \int_0^T u_t \delta B_t\right] = E\left[\int_0^T u_t D_t F dt\right]. \tag{6.12}$$

The next result, which is due to [30] and central in proving existence of strong solutions in the following, provides a compactness criterion for subsets of $L^2(\Omega)$ based on Malliavin calculus.

Proposition 6.6. *Let $F_n \in \mathbb{D}^{1,2}$, $n = 1, 2...$, be a given sequence of Malliavin differentiable random variables. Assume that there exist constants $\alpha > 0$ and $C > 0$ such that*

$$\sup_n E[|F_n|^2] \leq C,$$

$$\sup_n E\left[|D_t F_n - D_{t'} F_n|^2\right] \leq C|t - t'|^\alpha$$

for $0 \leq t' \leq t \leq T$, and

$$\sup_n \sup_{0\leq t\leq T} E\left[|D_t F_n|^2\right] \leq C\,.$$

Then the sequence F_n, $n = 1, 2...$, is relatively compact in $L^2(\Omega)$.

We conclude this review on Malliavin calculus by stating a relation between the Malliavin derivative and the first variation process of the solution of an SDE with smooth coefficients that is essential in the derivation of Theorem 6.1. We give the result for the case when the volatility coefficient is equal to 1, but the analogue result is valid for more general smooth volatility coefficients. Assume the drift coefficient $b(t, x)$ in the SDE (6.5) fulfils the Lipschitz and linear growth conditions (6.7)-(6.8). Then it is well known that there exists a unique strong solution X_t^x, $t \in [0, T]$, to equation (6.5) that is Malliavin differentiable, and that for all $0 \leq s \leq t \leq T$

the Malliavin derivative $D_s X_t^x$ fulfils, see e.g. [90, Theorem 2.2.1]

$$D_s X_t^x = 1 + \int_s^t b'(u, X_u^x) D_s X_u^x du, \tag{6.13}$$

where b' denotes the (weak) derivative of b with respect to x.

Further, under these assumptions the strong solution is also differentiable in its initial condition, and the first variation process $\frac{\partial}{\partial x} X_t^x$, $t \in [0,T]$, fulfils (see e.g. [70] for differentiable coefficients and [14] for an extension to Lipschitz coefficients)

$$\frac{\partial}{\partial x} X_t^x = 1 + \int_0^t b'(u, X_u^x) \frac{\partial}{\partial x} X_u^x du. \tag{6.14}$$

Solving equations (6.13) and (6.14) thus yields the following proposition.

Proposition 6.7. *Let X_t^x, $t \in [0,T]$, be the unique strong solution to equation (6.5) when $b(t,x)$ fulfils the Lipschitz and linear growth condition* (6.7)-(6.8). *Then X_t^x is Malliavin differentiable and differentiable in its initial condition for all $t \in [0,T]$, and for all $s \leq t \leq T$ we have*

$$D_s X_t^x = \exp\left\{ \int_s^t b'(u, X_u^x) du \right\} \tag{6.15}$$

and

$$\frac{\partial}{\partial x} X_t^x = \exp\left\{ \int_0^t b'(u, X_u^x) du \right\}. \tag{6.16}$$

As a consequence,

$$\frac{\partial}{\partial x} X_t^x = D_s X_t^x \frac{\partial}{\partial x} X_s^x, \tag{6.17}$$

where all equalities hold P-a.s.

6.2.2 Integration with respect to local-time

Let now X^x be a given (strong) solution to SDE (6.5). In the sequel we need the concept of stochastic integration over the plane with respect to the local time $L^{X^x}(t,y)$ of X^x. For Brownian motion, the local time integration theory in time and space has been introduced in [39]. We extend this local time integration theory to more general diffusions of type (6.5) by resorting to the Brownian setting under an equivalent measure where X^x is a Brownian motion. To this end, we notice the following fact that is extensively used throughout the paper.

Remark 6.8. *The Radon-Nikodym density*

$$\frac{dQ}{dP} = \mathcal{E}\left(- \int_0^T b(s, X_s^x) dB_s \right)$$

defines a probability measure Q equivalent to P under which X^x is Brownian motion starting in x. Indeed, because b is of at most linear growth we obtain by Grönwall's inequality as in the *proof of Lemma 6.29 a constant $C_{t,x} > 0$ such that $|X_t^x| \leq C_{t,x}(1 + |B_t|)$. One can thus find a equidistant partition $0 = t_0 < t_1 ... < t_m = T$ such that*

$$E\left[\exp\left\{\int_{t_i}^{t_{i+1}} b^2(s, X_s^x)ds\right\}\right] \leq E\left[\exp\left\{\int_{t_i}^{t_{i+1}} \left(C_1 + C_2|B_s| + C_3|B_s|^2\right) ds\right\} < \infty\right]$$

for all $i = 0, ..., m-1$, where C_1, C_2 and C_3 are some positive constants. Then it is well-known, see e.g. [62, Corollary 5.16], that Q is an equivalent probability measure under which X^x is Brownian motion by Girsanov's theorem.

We now define the feasible integrands for the local time-space integral with respect to $L^{X^x}(t, y)$ by the Banach space $(\mathcal{H}^x, \|\cdot\|)$ of functions $f : [0, T] \times \mathbb{R} \longrightarrow \mathbb{R}$ with norm

$$\begin{aligned}\|f\|_x = & 2\left(\int_0^T \int_{\mathbb{R}} f^2(s, y) \frac{1}{\sqrt{2\pi s}} \exp\left(-\frac{|y-x|^2}{2s}\right) dyds\right)^{1/2} \\ & + \int_0^T \int_{\mathbb{R}} |y - x| \, |f(s, y)| \, \frac{1}{s\sqrt{2\pi s}} \exp\left(-\frac{|y-x|^2}{2s}\right) dyds.\end{aligned}$$

We remark that this space of integrands is the same as the one introduced in [39] for Brownian motion (i.e. the special case when the X^x is a Brownian motion), except that we have in a straight forward manner generalised the space in [39] to the situation when the Brownian motion has arbitrary initial value x.

We denote by $f_\Delta : [0, T] \times \mathbb{R} \longrightarrow \mathbb{R}$ a simple function in the form

$$f_\Delta(s, y) = \sum_{1 \leq i \leq n-1, 1 \leq j \leq m-1} f_{ij} \mathbf{1}_{(y_i, y_{i+1}]}(y) \mathbf{1}_{(s_j, s_{j+1}]}(s),$$

where $(s_j)_{1 \leq j \leq m}$ is a partition of $[0, T]$ and $(y_i)_{1 \leq i \leq n}$ and $(f_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ are finite sequences of real numbers. It is readily checked that the space of simple functions is dense in $(\mathcal{H}^x, \|\cdot\|)$. The local time-space integral of an simple function f_Δ with respect to $L^{X^x}(dt, dy)$ is then defined by

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}} f_\Delta(s, y) L^{X^x}(ds, dy) := \\ &\qquad := \sum_{\substack{1 \leq i \leq n-1 \\ 1 \leq j \leq m-1}} f_{ij}(L^{X^x}(s_{j+1}, y_{i+1}) - L^{X^x}(s_j, y_{i+1}) - L^{X^x}(s_{j+1}, y_i) + L^{X^x}(s_j, y_i)).\end{aligned}$$

Lemma 6.9. *For $f \in \mathcal{H}^x$ let $f_n, n \geq 1$, be a sequence of simple functions converging to f in $\mathcal{H}^x$. Then $\int_0^T \int_{\mathbb{R}} f_n(s, y) L^{X^x}(ds, dy)$, $n \geq 1$, converges in probability. Further, for any other approximating sequence of simple functions the limit remains the same.*

Proof. Define $F_n^{X^x} := \int_0^T \int_{\mathbb{R}} f_n(s, x) L^{X^x}(ds, dx)$. Now consider the equivalent measure Q from Remark 6.8 under which X^x is Brownian motion. Define $F^{X^x} := \int_0^T \int_{\mathbb{R}} f(s, x) L^{X^x}(ds, dx)$ to be the time-space integral of f with respect to the local time of Brownian motion X^x under Q, which exists as an $L^1(Q)$-limit of $F_n^{X^x}, n \geq 1$ by the Brownian local time integration theory introduced in [39] (since $f_n, n \geq 1$ converge to f in $\mathcal{H}^x$). We show that $F_n^{X^x}, n \geq 1$ converge in probability to F^{X^x} under P. Indeed,

$$
\begin{aligned}
E[1 \wedge |F^{X^x} - F_n^{X^x}|] &= E\left[\left(1 \wedge |F^{B^x} - F_n^{B^x}|\right) \mathcal{E}\left(\int_0^T b(s, B_s^x) dB_s\right)\right] \\
&\leq E\left[\mathcal{E}\left(\int_0^T b(s, B_s^x) dB_s\right)^{1+\varepsilon}\right]^{1/(1+\varepsilon)} E\left[\left(1 \wedge |F^{B^x} - F_n^{B^x}|\right)^{\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{1+\varepsilon}} \\
&\leq C_\varepsilon E[\left(1 \wedge |F^{B^x} - F_n^{B^x}|\right)]^{\frac{\varepsilon}{1+\varepsilon}} \overset{n\to\infty}{\longrightarrow} 0\,,
\end{aligned} \tag{6.18}
$$

where, in analogy to the notation F^{X^x} and $F_n^{X^x}$ above, the notation F^{B^x} and $F_n^{B^x}$ refers to the corresponding integrals with respect to local time of Brownian motion B^x under P, and where in the first equality we have used that $(F^{B^x}, F_n^{B^x})$ has the same law under P as $(F^{X^x}, F_n^{X^x})$ under Q. The inequalities follow by Lemma 6.29 for some $\varepsilon > 0$ suitably small. Further, by [39] we know that $F_n^{B^x}, n \geq 1$ converge to F^{B^x} in $L^1(P)$, which implies the convergence in (6.18). Hence $F_n^{X^x}, n \geq 1$ converge to F^{X^x} in the Ky-Fan metric $d(X,Y) = E[1 \wedge |X - Y|]$, $X, Y \in L^0(\Omega)$, which characterises convergence in probability. Finally, again by [39], F^{X^x} is independent of the approximating sequence $f_n, n \geq 1$. □

Definition 6.10. *For $f \in \mathcal{H}^x$ the limit in Lemma 6.9 is called the time-space integral of f with respect to $L^{X^x}(dt, dx)$ and is denoted by $\int_0^T \int_{\mathbb{R}} f(s,y) L^{X^x}(ds, dy)$. Further, for any $t \in [0,T]$ we define $\int_0^t \int_{\mathbb{R}} f(s,y) L^{X^x}(ds, dy) := \int_0^T \int_{\mathbb{R}} f(s,y) I_{[0,t]}(s) L^{X^x}(ds, dy)$.*

Remark 6.11. *We notice that the drift coefficient $b(t,x)$ in (6.6), which is of linear growth in x uniformly in t, is in $\mathcal{H}^x$, and thus the local time integral of $b(t,x)$ with respect to $L^{X^x}(dt, dy)$ exists for any $x \in \mathbb{R}$.*

If X^x is a Brownian motion $B.$ we have the following decomposition due to [39] that we employ in the construction of strong solutions, and that also constitutes the foundation in the construction of the local time integral in [39].

Theorem 6.12. *Let $f \in \mathcal{H}^0$. Then*

$$
\begin{aligned}
&\int_0^t \int_{\mathbb{R}} f(s,y) L^{B^x}(ds, dy) = \\
&\qquad = -\int_0^t f(s, B_s^x) dB_s + \int_{T-t}^T f(T-s, \widehat{B}_s^x) dW_s - \int_{T-t}^T f(T-s, \widehat{B}_s^x) \frac{\widehat{B}_s}{T-s} ds,
\end{aligned} \tag{6.19}
$$

where $\widehat{B}_t = B_{T-t}, 0 \leq t \leq T$ is time-reversed Brownian motion, and $W.$, defined by

$$
\widehat{B}_t = B_T + W_t - \int_0^t \frac{\widehat{B}_s}{T-s} ds,
$$

is a Brownian motion with respect to the filtration of $\widehat{B}.$.

We conclude this subsection by stating three further identities for the local time integral of a general diffusions X^x which will be useful later on.

Lemma 6.13. *Let $f \in \mathcal{H}^x$ be Lipschitz continuous in x. Then for all $t \in [0,T]$*

$$-\int_0^t \int_{\mathbb{R}} f(s,y) L^{X^x}(ds,dy) = \int_0^t f'(s,X_s^x)ds. \tag{6.20}$$

where f' denotes the (weak) derivative of $f(t,y)$ with respect to y.

If $f \in \mathcal{H}^x$ is time homogeneous (i.e. $f(t,y) = f(y)$ only depends on the space variable) and locally square integrable, then for any $t \in [0,T]$

$$\int_0^t \int_{\mathbb{R}} f(s,x) L^{X^x}(ds,dx) = -[f(\cdot,X^x),X^x]_t. \tag{6.21}$$

and

$$-\int_0^t \int_{\mathbb{R}} f(s,y) L^{X^x}(ds,dy) = 2F(X_t^x) - 2F(x) - 2\int_0^t f(X_s^x)dX_s^x \tag{6.22}$$

where F is a primitive function of f and $[\tilde{b}(\cdot,X_\cdot^x),X_\cdot^x]_t$ is the generalised covariation process

$$[f(\cdot,X_\cdot^x),X_\cdot^x]_t := P - \lim_{m\to\infty} \sum_{k=1}^m \left(f(t_k^m, X_{t_k^m}^x) - f(t_{k-1}^m, X_{t_{k-1}^m}^x) \right) \left(X_{t_k}^x - X_{t_{k-1}}^x \right)$$

where for every m $\{t_k^m\}_{k=1}^m$ is a partition of $[0,t]$ such that $\lim_m \sup_{k=1,\dots,m} |t_k^m - t_{k-1}^m| = 0$. Note that (6.22) *can be considered as a generalised Itô formula.*

Proof. If $X^x = x + B$, then identities (6.20)-(6.22) are given in [39]. For general X^x, we consider the identities under the equivalent measure Q from Remark 6.8. Then, by the construction of the local time integral outlined in Lemma 6.9, the integrals in the identities are the ones with respect to Brownian motion X^x, for which we know the identities hold by [39] (where such identities are given in the case $x = 0$ but one can easily extend them to the case of the Brownian motion starting at an arbitrary $x \in \mathbb{R}$). □

6.3 Existence, Malliavin and Sobolev differentiability of strong solutions

In this section we prepare the necessary theoretical grounds to develop the probabilistic representation of Deltas. Being notationally and technically rather heavy, the proofs of this section are deferred to Appendix 6.A for an improved flow and readability of the paper. We first study the existence and Malliavin differentiability of a unique strong solution of SDE (6.5) before we turn to the differentiability of the strong solution in its initial condition and the corresponding first variation process. We state the first main result of this section:

Theorem 6.14. *Suppose that the drift coefficient $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ is in the form (6.6). Then there exists a unique strong solution $\{X_t^x\}_{t\in[0,T]}$ to SDE (6.5). In addition, X_t^x is Malliavin differentiable for every $t \in [0,T]$.*

The proof of Theorem 6.14 employs several auxiliary results presented in Appendix 6.A. The main steps are:

1. First, we construct a weak solution X^x to (6.5) by means of Girsanov's theorem, that is we introduce a probability space $(\Omega, \mathcal{F}, P)$ that carries some Brownian motion B and a continuous process X^x such that (6.5) is fulfilled. However, a priori X^x is not adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ generated by Brownian motion B.

2. Next, we approximate the drift coefficient $b = \tilde{b} + \hat{b}$ by a sequence of functions (which always exists by standard approximation results)

$$b_n := \tilde{b}_n + \hat{b}, \quad n \geq 1, \tag{6.23}$$

   such that $\{\tilde{b}_n\}_{n\geq 1} \subset C_0^\infty([0,T]\times\mathbb{R})$ with $\sup_{n\geq 1} \|\tilde{b}_n\|_\infty \leq C < \infty$ and $\tilde{b}_n \to \tilde{b}$ in $(t,x) \in [0,T] \times \mathbb{R}$ a.e. with respect to the Lebesgue measure. By standard results on SDE's, we know that for each smooth coefficient b_n, $n \geq 1$, there exists a unique strong solution $X_\cdot^{n,x}$ to the SDE

$$dX_t^{n,x} = b_n(t, X_t^{n,x})dt + dB_t, \quad 0 \leq t \leq T, \quad X_0^{n,x} = x \in \mathbb{R}\,. \tag{6.24}$$

   We then show that for each $t \in [0,T]$ the sequence $X_t^{n,x}$ converges weakly to the conditional expectation $E[X_t^x|\mathcal{F}_t]$ in the space $L^2(\Omega;\mathcal{F}_t)$ of square integrable, $\mathcal{F}_t$-measurable random variables.

3. By Proposition 6.7 we know that for each $t \in [0,T]$ the strong solutions $X_t^{n,x}$, $n \geq 1$, are Malliavin differentiable with

$$D_s X_t^{n,x} = \exp\left\{\int_s^t b_n'(u, X_u^{n,x})du\right\}, \quad 0 \leq s \leq t \leq T, \quad n \geq 1, \tag{6.25}$$

   where b_n' denotes the derivative of b_n with respect to x. We will use representation (6.25) to employ a compactness criterion based on Malliavin calculus to show that for every $t \in [0,T]$ the set of random variables $\{X_t^{n,x}\}_{n\geq 1}$ is relatively compact in $L^2(\Omega;\mathcal{F}_t)$, which then allows to conclude that $X_t^{n,x}$ converges strongly in $L^2(\Omega;\mathcal{F}_t)$ to $E[X_t^x|\mathcal{F}_t]$. Further we obtain that $E[X_t^x|\mathcal{F}_t]$ is Malliavin differentiable as a consequence of the compactness criterion.

4. In the last step we show that $E[X_t^x|\mathcal{F}_t] = X_t^x$, which implies that X_t^x is $\mathcal{F}_t$-measurable and thus a strong solution. Moreover, we show that this solution is unique.

Notation: In the following we sometimes include the drift coefficient b into the sequence $\{b_n\}_{n\geq 0}$ by putting $b_0 := \tilde{b}_0 + \hat{b} := \tilde{b} + \hat{b} = b$.

The next important result is an explicit representation of the Malliavin derivative of the strong solution X_t^x, $t \in [0,T]$. For smooth coefficients b we can explicitly express the Malliavin derivative in terms of the derivative of b as stated in (6.25). For general, not necessarily differentiable coefficients b, we are still able to give an explicit formula which now only involves the coefficient b in a local time integral:

Proposition 6.15. *For $0 \leq s \leq t \leq T$, the Malliavin derivative $D_s X_t^x$ of the unique strong solution X_t^x to equation (6.5) has the following explicit representation:*

$$D_s X_t^x = \exp\left\{ - \int_s^t \int_{\mathbb{R}} b(u,y) L^{X^x}(du, dy) \right\} \text{ } P\text{-a.s.}, \tag{6.26}$$

where $L^{X^x}(du, dy)$ denotes integration in space and time with respect to the local time of X^x, see Section 6.2.2 for definitions.

Next, we turn our attention to the study of the strong solution X_t^x as a function in its initial condition x for SDE's with possible irregular drift coefficients. The first result establishes Hölder continuity jointly in time and space.

Proposition 6.16. *Let X_t^x, $t \in [0, T]$ be the unique strong solution to the SDE* (6.5)*. Then for all $s, t \in [0, T]$ and $x, y \in K$ for any arbitrary compact subset $K \subset \mathbb{R}$ there exists a constant $C = C(K, \|\tilde{b}\|_\infty, \|\hat{b}'\|_\infty) > 0$ such that*

$$E\left[|X_t^x - X_s^y|^2 \right] \leq C(|t-s| + |x-y|^2).$$

In particular, there exists a continuous version of the random field $(t, x) \mapsto X_t^x$ with Hölder continuous trajectories of order $\alpha < 1/2$ in $t \in [0, T]$ and $\alpha < 1$ in $x \in \mathbb{R}$.

If the drift coefficient b is regular, then we know by Proposition 6.7 that X_t^x is even differentiable as a function in x. The first variation process $\frac{\partial}{\partial x} X_\cdot^x$ is then given by (6.16) in terms of the derivative of the drift coefficient and is closely related to the Malliavin derivative by (6.17). In the following we will derive analogous results for irregular drift coefficients, where in general the first variation process will now exist in the Sobolev derivative sense. Let $U \subset \mathbb{R}$ be an open and bounded subset. The Sobolev space $W^{1,2}(U)$ is defined as the set of functions $u : \mathbb{R} \to \mathbb{R}$, $u \in L^2(U)$ such that its weak derivative belongs to $L^2(U)$. We endow this space with the norm

$$\|u\|_{1,2} = \|u\|_2 + \|u'\|_2$$

where u' stands for the *weak derivative* of $u \in W^{1,2}(U)$. We say that the solution X_t^x, $t \in [0, T]$, is *Sobolev differentiable* in U if for all $t \in [0, T]$, $X_t^\cdot$ belongs to $W^{1,2}(U)$, P-a.s. Observe that in general $X_t^\cdot$ is not in $W^{1,2}(\mathbb{R})$, e.g. take $b \equiv 0$.

Theorem 6.17. *Let $b : [0, T] \times \mathbb{R} \to \mathbb{R}$ be as in (6.6). Let X_t^x , $t \in [0, T]$ be the unique strong solution to the SDE (6.5) and $U \subset \mathbb{R}$ an open, bounded set. Then for every $t \in [0, T]$ we have*

$$(x \mapsto X_t^x) \in L^2(\Omega, W^{1,2}(U)).$$

We remark that using analogue techniques as in [87] one could even establish that the strong solution gives rise to a flow of Sobolev diffeomorphisms. This, however, is beyond the scope of this paper.

Similarly as for the Malliavin derivative, we are able to give an explicit representation for the first variation process in the Sobolev sense that does not involve the derivative of the drift coefficient by employing local time integration.

Proposition 6.18. *Let $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ be as in (6.6). Then the first variation process (in the Sobolev sense) of the strong solution X_t^x, $t \in [0,T]$ to SDE (6.5) has the following explicit representation*

$$\frac{\partial}{\partial x} X_t^x = \exp\left\{ -\int_0^t \int_{\mathbb{R}} b(u,y) L^{X^x}(du,dy) \right\} \quad dt \otimes P - a.s. \tag{6.27}$$

As a direct consequence of Proposition 6.18 together with Proposition 6.15 we obtain the following relation between the Malliavin derivative and the first variation process, which is an extension of Proposition 6.7 to irregular drift coefficients and which is a key result in deriving the desired expression for the Delta.

Corollary 6.19. *Let X_t^x, $t \in [0,T]$, be the unique strong solution to (6.5). Then the following relationship between the spatial derivative and the Malliavin derivative of X_t^x holds:*

$$\frac{\partial}{\partial x} X_t^x = D_s X_t^x \frac{\partial}{\partial x} X_s^x \;\; P - a.s. \tag{6.28}$$

for any $s, t \in [0,T]$, $s \le t$.

Remark 6.20. *Note that by Lemma 6.13 the Malliavin derivative in* (6.26) *and the first variation process in* (6.27) *can be expressed in various alternative ways. Firstly, we observe that by formula (6.20) the local time integral of the regular part $\hat{b}$ in b can be separated and rewritten in the form*

$$-\int_s^t \int_{\mathbb{R}} b(u,y) L^{X^x}(du,dy) = -\int_s^t \int_{\mathbb{R}} \tilde{b}(u,y) L^{X^x}(du,dy) + \int_s^t \hat{b}'(u, X_u^x) du \text{ a.s.} \tag{6.29}$$

If in addition $\tilde{b}(t,\cdot)$ is locally square integrable and continuous in t as a map from $[0,T]$ to $L^2_{loc}(\mathbb{R})$ or even time-homogeneous, then by Lemma 6.13 also the local time integral associated to the irregular part $\tilde{b}$ can be reformulated in terms of the generalised covariation process as in (6.21) *or in terms of the generalised Itô formula as in* (6.22)*, respectively. In particular, these reformulations appear to be useful for simulation purposes.*

We conclude this section by giving an extension of all the results seen so far to a class of SDE's with more general diffusion coefficients.

Theorem 6.21. *Consider the time-homogeneous SDE*

$$dX_t^x = b(X_t^x)dt + \sigma(X_t^x)dB_t, \quad X_0^x = x \in \mathbb{R}, \quad 0 \le t \le T, \tag{6.30}$$

where the coefficients $b : \mathbb{R} \longrightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \longrightarrow \mathbb{R}$ are Borel measurable. Require that there exists a twice continuously differentiable bijection $\Lambda : \mathbb{R} \longrightarrow \mathbb{R}$ with derivatives Λ' and Λ'' such that

$$\Lambda'(y)\sigma(y) = 1 \text{ for a.e. } y \in \mathbb{R},$$

as well as

$$\Lambda^{-1} \text{ is Lipschitz continuous.}$$

Suppose that the function $b_* : \mathbb{R} \longrightarrow \mathbb{R}$ *given by*

$$b_*(x) := \Lambda'\left(\Lambda^{-1}(x)\right) b(\Lambda^{-1}(x)) + \frac{1}{2}\Lambda''\left(\Lambda^{-1}(x)\right) \sigma(\Lambda^{-1}(x))^2$$

satisfies the conditions of Theorem 6.14. Then there exists a Malliavin differentiable strong solution $X_\cdot^x$ *to (6.30) which is (locally) Sobolev differentiable in its initial condition.*

Proof. The proof is obtained directly from Itô's formula. See [83]. □

6.4 Existence and derivative-free representations of the Delta

We now turn the attention to the study of option price sensitivities with respect to the initial value of an underlying asset with irregular drift coefficient. Notably, we will consider lookback options with path-dependent (discounted) pay-off in the form

$$\Phi(X_{T_1}^x, \ldots, X_{T_m}^x) \tag{6.31}$$

for time points $T_1, \ldots, T_m \in (0, T]$, some function $\Phi : \mathbb{R}^m \to \mathbb{R}$, and where the evolution of the underlying price process under the risk-neutral pricing measure is modelled by the strong solution X^x of SDE (6.5) with possibly irregular drift b as in (6.6). In particular, for $m = 1$ the pay-off (6.31) is associated to a European option with maturity T_1. Then the current option price is given by $E\left[\Phi(X_{T_1}^x, \ldots, X_{T_m}^x)\right]$ and the main objective of this section is to establish existence and a derivative-free, probabilistic representation of the Delta

$$\frac{\partial}{\partial x} E\left[\Phi(X_{T_1}^x, \ldots, X_{T_m}^x)\right] .$$

After having analysed lookback options, we will also address the problem of computing Deltas of Asian options with (discounted) path-dependent pay-off in the form

$$\Phi\left(\int_{T_1}^{T_2} X_u^x \, du\right) \tag{6.32}$$

for $T_1, T_2 \in [0, T]$ and some function $\Phi : \mathbb{R} \to \mathbb{R}$.

We start with a preliminary result which shows that in case of a smooth pay-off function with compact support the Delta exists for a large class of path dependent options that includes both lookback as well as Asian options.

Lemma 6.22. *Let* X_t^x*,* $t \in [0, T]$*, be the strong solution to SDE (6.5) and* $\{X_t^{n,x}\}_{n \geq 1}$ *the corresponding approximating strong solutions of SDE (6.24). Let* $\Phi \in C_0^\infty(\mathbb{R}^m)$ *and consider the functions*

$$u_n(x) := E\left[\Phi\left(\int_0^T X_u^{n,x} \mu_1(du), \int_0^T X_u^{n,x} \mu_2(du), \ldots, \int_0^T X_u^{n,x} \mu_m(du)\right)\right] \tag{6.33}$$

and

$$u(x) := E\left[\Phi\left(\int_0^T X_u^x \mu_1(du), \int_0^T X_u^x \mu_2(du), \ldots, \int_0^T X_u^x \mu_m(du)\right)\right] \tag{6.34}$$

where $\mu_1, \ldots, \mu_m$ *are finite measures on* $[0,T]$ *independent of* $x \in \mathbb{R}$. *Consider also the function*

$$\bar{u}(x) := E\left[\sum_{i=1}^m \partial_i \Phi\left(\int_0^T X_u^x \mu_1(du), \int_0^T X_u^x \mu_2(du), \ldots, \int_0^T X_u^x \mu_m(du)\right) \int_0^T \frac{\partial}{\partial x} X_u^x \mu_i(du)\right] \tag{6.35}$$

where $\frac{\partial}{\partial x} X^x$ *is the first variation process of* X^x *introduced in* (6.27). *Then*

$$u_n(x) \xrightarrow{n\to\infty} u(x) \quad \textit{for all } \ x \in \mathbb{R},$$

and

$$u_n'(x) \xrightarrow{n\to\infty} \bar{u}(x)$$

uniformly on compact subsets $K \subset \mathbb{R}$, *where* u_n' *denotes the derivative. As a result, we obtain that* $u \in C^1(\mathbb{R})$ *with* $u' = \bar{u}$. *In particular, we obtain the result for lookback options by taking* $\mu_i = \delta_{t_i}$ *the Dirac measure concentrated on* t_i, $i = 1, \ldots, m$, *and for Asian options by taking* $m = 1$ *and* $\mu_1 = du$.

Proof. First of all, observe that the expression in (6.35) is well-defined. This can be seen by using Cauchy-Schwarz inequality, the fact that $\Phi \in C_0^\infty(\mathbb{R}^m)$, and Corollary 6.37.

It is readily checked that $u_n(x) \to u(x)$ for all $x \in \mathbb{R}$ since Φ is smooth by using the mean-value theorem and the fact that $X_t^{n,x} \to X_t^x$ in $L^2(\Omega)$ as $n \to \infty$ for every $t \in [0,T]$ (see Theorem 6.34).

We introduce the following short-hand notation for the m-dimensional random vector associated to a process Y:

$$h(Y_{\cdot,T}) := \left(\int_0^T Y_u \mu_1(du)\ , \ \int_0^T Y_u \mu_2(du)\ , \ \ldots\ , \ \int_0^T Y_u \mu_m(du)\right).$$

For the smooth coefficients b_n we have $u_n \in C^1(\mathbb{R})$, $n \geq 1$, and since $\partial_i \Phi$ are bounded for all $i = 1, \ldots, m$ and by dominated convergence we have

$$u_n'(x) = E\left[\sum_{i=1}^m \partial_i \Phi\left(h(X_{\cdot,T}^{n,x})\right) \int_0^T \frac{\partial}{\partial x} X_u^{n,x} \mu_i(du)\right].$$

Moreover, we can take integration with respect to $\mu_i(du)$, $i = 1, ...m$, outside the expectation. Thus

$$u_n'(x) = \sum_{i=1}^m \int_0^T E\left[\partial_i \Phi\left(h(X_{\cdot,T}^{n,x})\right) \frac{\partial}{\partial x} X_u^{n,x}\right] \mu_i(du).$$

Hence

$$\begin{aligned}|u_n'(x)-\bar{u}(x)| &= \sum_{i=1}^m \int_0^T E\left[\partial_i\Phi\left(h(X_{\cdot,T}^{n,x})\right)\frac{\partial}{\partial x}X_u^{n,x} - \partial_i\Phi\left(h(X_{\cdot,T}^{x})\right)\frac{\partial}{\partial x}X_u^{x}\right]\mu_i(du)\\ &=: \sum_{i=1}^m \int_0^T F_{n,i}(u,x)\mu_i(du)\end{aligned}$$

where $F_{n,i}(u,x)$ denotes the expectation in the integral. We will show that for any $i=1,\dots,m$ and compact subset $K\subset\mathbb{R}$,

$$\sup_{(u,x)\in[0,T]\times K}|F_{n,i}(u,x)| \xrightarrow{n\to\infty} 0.$$

Indeed, by plugging in expression (6.27) for the first variation process and Girsanov's theorem we get

$$\begin{aligned}|F_{n,i}(u,x)| \leq &\Bigg|E\Big[\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)\exp\left\{-\int_0^u\int_{\mathbb{R}} b_n(v,y)L^{B^x}(dv,dy)\right\}\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)\\ &-\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)\exp\left\{-\int_0^u\int_{\mathbb{R}} b(v,y)L^{B^x}(dv,dy)\right\}\mathcal{E}\left(\int_0^T b(u,B_u^x)dB_u\right)\Big]\Bigg|\\ \leq &\Bigg|E\Big[\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)\mathcal{E}\left(\int_0^T b(u,B_u^x)dB_u\right)\\ &\times\left(\exp\left\{-\int_0^u\int_{\mathbb{R}} b_n(v,y)L^{B^x}(dv,dy)\right\}-\exp\left\{-\int_0^u\int_{\mathbb{R}} b(v,y)L^{B^x}(dv,dy)\right\}\right)\Big]\Bigg|\\ &+\Bigg|E\Big[\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)\exp\left\{-\int_0^u\int_{\mathbb{R}} b_n(v,y)L^{B^x}(dv,dy)\right\}\\ &\times\left(\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)-\mathcal{E}\left(\int_0^T b(u,B_u^x)dB_u\right)\right)\Big]\Bigg|\\ :=&\, I_n+II_n\end{aligned}$$

Here, we will show estimates for II_n, for I_n the argument is analogous. Similarly to how we obtain the estimate $II_n^1+II_n^2$ in the proof of Lemma 6.33, using inequality $|e^x-1|\leq|x|(e^x+1)$ we get

$$\begin{aligned}II_n \lesssim\, & E\Big[|\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)||U_n|\exp\left\{-\int_0^u\int_{\mathbb{R}} b_n(v,y)L^{B^x}(dv,dy)\right\}\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)\Big]\\ &+E\Big[|\partial_i\Phi\left(h(B_{\cdot,T}^x)\right)||U_n|\exp\left\{-\int_0^u\int_{\mathbb{R}} b_n(v,y)L^{B^x}(dv,dy)\right\}\mathcal{E}\left(\int_0^T b(u,B_u^x)dB_u\right)\Big]\\ =:&\, II_n^1+II_n^2,\end{aligned}$$

where

$$U_n := \int_0^T (\tilde{b}_n(u, B_u^x) - \tilde{b}(u, B_u^x))dB_u - \frac{1}{2}\int_0^T (b_n^2(u, B_u^x) - b^2(u, B_u^x))du.$$

We will now show that $II_n^1 \to 0$ as $n \to \infty$ uniformly in x on a compact subset $K \subset \mathbb{R}$. The convergence of II_n^2 then follows immediately, too. Denote $p = \frac{1+\varepsilon}{\varepsilon}$ with $\varepsilon > 0$ from Lemma 6.29 and use Hölder's inequality with exponent $1+\varepsilon$ on the Doléans-Dade exponential, then employ formula (6.20) on $\hat{b}$ in $b_n = \tilde{b}_n + \hat{b}$ and use Cauchy-Schwarz inequality successively. As a result,

$$\begin{aligned} II_n^1 \lesssim E\left[\mathcal{E}\left(\int_0^T b_n(u, B_u^x)dB_u\right)^{1+\varepsilon}\right]^{1/(1+\varepsilon)} & E\left[|\partial_i \Phi\left(h(B_{\cdot,T^*}^x)\right)|\right]^{1/(2p)} E[|U_n|^{8p}]^{1/(8p)} \\ \times E\left[\exp\left\{-4p\int_0^u \int_{\mathbb{R}} \tilde{b}_n(v,y) L^{B^x}(dv,dy)\right]^{1/(4p)}\right\} & E\left[\exp\left\{8p\int_0^u \hat{b}'(v, B_v^x)dv\right\}\right]^{1/(8p)}. \end{aligned}$$

The first and fourth factor are bounded uniformly in $n \geq 0$ and $x \in K$ by Remark 6.30 and Lemma 6.31, respectively. The second and and fifth factor can be controlled since $\partial_i \Phi$, $i = 1, \ldots, m$ and $\hat{b}'$ are uniformly bounded. It remains to show that

$$\sup_{x \in K} E[|U_n|^{8p}] \xrightarrow{n \to \infty} 0$$

for any compact subset $K \subset \mathbb{R}$.

Using Minkowski's inequality, Burkholder-Davis-Gundy's inequality and Hölder's inequality we can write

$$E[|U_n|^{8p}] \lesssim \int_0^T E[|\tilde{b}_n(u, B_u^x) - \tilde{b}(u, B_u^x)|^{8p}]du + \int_0^T E[|b_n^2(u, B_u^x) - b^2(u, B_u^x)|^{8p}]du. \tag{6.36}$$

Then write the integrand of the first term in (6.36) as

$$E[|\tilde{b}_n(u, B_u^x) - \tilde{b}(u, B_u^x)|^{8p}] = \frac{1}{\sqrt{2\pi u}}\int_{\mathbb{R}} |\tilde{b}_n(u,y) - \tilde{b}(u,y)|^{8p} e^{-\frac{(y-x)^2}{2u}} dy.$$

Using Cauchy-Schwarz inequality on $|\tilde{b}_n(u,y) - \tilde{b}(u,y)|^{8p} e^{\frac{-y^2}{4u}}$ we obtain

$$\begin{aligned} E[|\tilde{b}_n(u, B_u^x) - & \tilde{b}(u, B_u^x)|^{8p}] \leq \\ & \leq \frac{1}{\sqrt{2\pi u}} e^{-\frac{x^2}{2u}} \left(\int_{\mathbb{R}} |\tilde{b}_n(u,y) - \tilde{b}(u,y)|^{16p} e^{-\frac{y^2}{2u}} dy\right)^{1/2} \left(\int_{\mathbb{R}} e^{-\frac{y^2}{2u} + 2\frac{xy}{u}} dy\right)^{1/2}. \end{aligned}$$

Then for each $u \in [0,T]$, by taking the supremum over $x \in K$ and by dominated convergence, we get

$$\sup_{x \in K} E[|\tilde{b}_n(u, B_u^x) - \tilde{b}(u, B_u^x)|^{8p}] \xrightarrow{n \to \infty} 0\,,$$

and hence the result follows. Similarly, one can argue for the second term in (6.36).

In sum,

$$\sup_{(u,x)\in[0,t]\times K} |F_{n,i}(u,x)| \xrightarrow{n\to\infty} 0$$

for all $i = 1,\dots,m$ and hence $u_n'(x) \xrightarrow{n\to\infty} \bar{u}(x)$ uniformly on compact sets $K \subset \mathbb{R}$, and thus $u \in C^1(\mathbb{R})$ with $u' = \bar{u}$. □

We come to the main result of this paper, which extends Theorem 6.1 to lookback options written on underlyings with irregular drift coefficients. In particular, when plugging in expression (6.27) for the first variation process, we see that the formula for the Delta in (6.38) below involves neither the derivative of the pay-off function Φ nor the derivative of the drift coefficient b. We obtain this result for pay-off functions $\Phi \in L_w^q(\mathbb{R}^m)$, where

$$L_w^q(\mathbb{R}^m) := \left\{ f : \mathbb{R}^m \to \mathbb{R} \text{ measurable} : \int_{\mathbb{R}^m} |f(x)|^q \, w(x)dx < \infty \right\} \tag{6.37}$$

for the weight function w defined by $w(x) := \exp(-\frac{1}{2T}|x|^2)$, $x \in \mathbb{R}^m$, and where the exponent q depends on the drift b. Note that all pay-off functions of practical relevance are contained in these spaces.

Theorem 6.23. *Let X^x be the strong solution to SDE* (6.5) *and $\Phi : \mathbb{R}^m \to \mathbb{R}$ a function in $L_w^{4p}(\mathbb{R}^m)$, where $p > 1$ is the conjugate of $1+\varepsilon$ for $\varepsilon > 0$ in Lemma 6.29. Then, for any $0 < T_1 \le \cdots \le T_m \le T$, the price*

$$u(x) := E\left[\Phi(X_{T_1}^x,\dots,X_{T_m}^x)\right]$$

of the associated lookback option is continuously differentiable in $x \in \mathbb{R}$, and its derivative, i.e. the Delta, takes the form

$$u'(x) = E\left[\Phi(X_{T_1}^x,\dots,X_{T_m}^x)\int_0^T a(s)\frac{\partial}{\partial x}X_s^x \, dB_s\right] \tag{6.38}$$

for any bounded measurable function $a : \mathbb{R} \to \mathbb{R}$ such that, for every $i = 1,\dots,m$,

$$\int_0^{T_i} a(s)ds = 1. \tag{6.39}$$

Proof. Assume first $\Phi \in C_0^\infty(\mathbb{R}^m)$. Then by Lemma 6.22 with $\mu_i = \delta_{t_i}$, $i = 1,\dots,m$, we know that $u(x) = E\left[\Phi(X_{T_1}^x,\dots,X_{T_m}^x)\right]$ is continuously differentiable with derivative

$$u'(x) := \sum_{i=1}^m E\left[\partial_i\Phi(X_{T_1}^x,\dots,X_{T_m}^x)\frac{\partial}{\partial x}X_{T_i}^x\right].$$

Now, by Corollary 6.19, we have for any $i = 1,\dots,m$

$$\frac{\partial}{\partial x}X_{T_i}^x = D_s X_{T_i}^x \frac{\partial}{\partial x}X_s^x \text{ for all } s \le T_i\,. \tag{6.40}$$

Also recall that $D_s X_{T_i}^x = 0$ for $s \ge T_i$. So, for any function $a : \mathbb{R} \to \mathbb{R}$ satisfying (6.39) we have

$$\frac{\partial}{\partial x}X_{T_i}^x = \int_0^T a(s)D_s X_{T_i}^x \frac{\partial}{\partial x}X_s^x ds.$$

As a result,

$$\begin{aligned}
u'(x) &= \sum_{i=1}^m E\left[\partial_i \Phi(X_{T_1}^x, \dots, X_{T_m}^x)\int_0^T a(s)D_s X_{T_i}^x \frac{\partial}{\partial x}X_s^x ds\right] \\
&= E\left[\int_0^T a(s)D_s\Phi(X_{T_1}^x, \dots, X_{T_m}^x)\frac{\partial}{\partial x}X_s^x ds\right],
\end{aligned}$$

where in the last step we could use the chain rule for the Malliavin derivative backwards, see Lemma 6.4, since $\Phi(X_{T_1}^x, \dots, X_{T_m}^x)$ is Malliavin differentiable due to Theorem 6.14. Then $a(s)\frac{\partial}{\partial x}X_s^x$ is an $\mathcal{F}_s$-adapted Skorokhod integrable process by Corollary 6.37 with $p = 2$, so the duality formula for the Malliavin derivative (see Theorem 6.5) yields

$$u'(x) = E\left[\Phi(X_{T_1}^x, \dots, X_{T_m}^x)\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s\right].$$

Finally, we extend the result to a pay-off function $\Phi \in L_w^{4p}(\mathbb{R}^m)$. By standard arguments we can approximate Φ by a sequence of functions $\Phi_n \in C_0^\infty(\mathbb{R}^m)$, $n \geq 0$, such that $\Phi_n \to \Phi$ in $L_w^{4p}(\mathbb{R}^m)$ as $n \to \infty$. Now define

$$u_n(x) := E[\Phi_n(X_{T_1}^x, \dots, X_{T_m}^x)]$$

and

$$\bar{u}(x) := E[\Phi(X_{T_1}^x, \dots, X_{T_m}^x)\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s].$$

Then

$$\begin{aligned}
|u_n'(x) - \bar{u}(x)| &= \left|E\left[\left(\Phi_n(X_{T_1}^x, \dots, X_{T_m}^x) - \Phi(X_{T_1}^x, \dots, X_{T_m}^x)\right)\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s\right]\right| \\
&\leq E\left[\left|\Phi_n(X_{T_1}^x, \dots, X_{T_m}^x) - \Phi(X_{T_1}^x, \dots, X_{T_m}^x)\right|^2\right]^{1/2} E\left[\int_0^T |a(s)\frac{\partial}{\partial x}X_s^x|^2 ds\right]^{1/2} \\
&\leq CE\left[\left|\Phi_n(B_{T_1}^x, \dots, B_{T_m}^x) - \Phi(B_{T_1}^x, \dots, B_{T_m}^x)\right|^2 \mathcal{E}\left(\int_0^T b(u, B_u^x)dB_u\right)\right]^{1/2},
\end{aligned}$$

where we have used Cauchy-Schwarz inequality, Itô's isometry, Corollary 6.37 and Girsanov's theorem in this order. Then we apply Hölder's inequality with $1 + \varepsilon$ for a small enough $\varepsilon > 0$ and use Lemma 6.29 to get

$$\begin{aligned}
&|u_n'(x) - \bar{u}(x)| \leq \\
&\leq CE\left[\left|\Phi_n(B_{T_1}^x, \dots, B_{T_m}^x) - \Phi(B_{T_1}^x, \dots, B_{T_m}^x)\right|^{2\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} E\left[\mathcal{E}\left(\int_0^T b(u, B_u^x)dB_u\right)^{1+\varepsilon}\right]^{\frac{1}{2(1+\varepsilon)}} \\
&\leq CE\left[\left|\Phi_n(B_{T_1}^x, \dots, B_{T_m}^x) - \Phi(B_{T_1}^x, \dots, B_{T_m}^x)\right|^{2\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}}.
\end{aligned}$$

For the last quantity, denote by $P_t(y) := \frac{1}{\sqrt{2\pi t}} e^{-y^2/(2t)}$, $y \in \mathbb{R}$ the density of B_t, and set $T_0 := 0$ and $y_0 := x$. Recall that $0 < T_1 \leq \cdots \leq T_m$. Using the independent increments of the Brownian motion we rewrite

$$E\left[\left|\Phi_n(B^x_{T_1}, \ldots, B^x_{T_m}) - \Phi(B^x_{T_1}, \ldots, B^x_{T_m})\right|^{2\frac{1+\varepsilon}{\varepsilon}} \right]$$
$$= \int_{\mathbb{R}^m} |\Phi_n(y_1, \ldots, y_m) - \Phi(y_1, \ldots, y_m)|^{2\frac{1+\varepsilon}{\varepsilon}} \prod_{i=1}^m P_{T_i - T_{i-1}}(y_i - y_{i-1}) dy_1 \cdots dy_m.$$

Furthermore, with $t^* := \min_{i=1,\ldots,m-1}(t_{i+1} - t_i)$

$$E\left[\left|\Phi_n(B^x_{T_1}, \ldots, B^x_{T_m}) - \Phi(B^x_{T_1}, \ldots, B^x_{T_m})\right|^{2\frac{1+\varepsilon}{\varepsilon}} \right]$$
$$\leq (2\pi t^*)^{-m/2} \int_{\mathbb{R}^m} |\Phi_n(y_1, \ldots, y_m) - \Phi(y_1, \ldots, y_m)|^{2\frac{1+\varepsilon}{\varepsilon}} \prod_{i=1}^m e^{-\frac{y_i^2}{4(T_i - T_{i-1})}}$$
$$\times e^{-\frac{y_i^2}{4(T_i - T_{i-1})} + \frac{y_i y_{i-1}}{T_i - T_{i-1}} - \frac{y_{i-1}^2}{2(T_i - T_{i-1})}} dy_1 \cdots dy_m.$$

By applying Cauchy-Schwarz inequality we obtain

$$E\left[\left|\Phi_n(B^x_{T_1}, \ldots, B^x_{T_m}) - \Phi(B^x_{T_1}, \ldots, B^x_{T_m})\right|^{2\frac{1+\varepsilon}{\varepsilon}} \right]$$
$$\leq (2\pi t^*)^{-m/2} \left(\int_{\mathbb{R}^m} |\Phi_n(y_1, \ldots, y_m) - \Phi(y_1, \ldots, y_m)|^{4\frac{1+\varepsilon}{\varepsilon}} e^{-\frac{|y|^2}{2T}} dy_1 \cdots dy_m \right)^{1/2}$$
$$\times \left(\int_{\mathbb{R}^m} \prod_{i=1}^m e^{-\frac{y_i^2}{2(T_i - T_{i-1})} + \frac{2 y_i y_{i-1}}{T_i - T_{i-1}} - \frac{y_{i-1}^2}{(T_i - T_{i-1})}} dy_1 \cdots dy_m \right)^{1/2}$$
$$=: I_n \cdot II.$$

For the second factor we have

$$II \leq e^{-\frac{x^2}{T}} \left(\int_{\mathbb{R}^m} e^{-\frac{y_1}{2T} + \frac{x y_1}{T}} \prod_{i=2}^m e^{-\frac{(y_i - y_{i-1})^2}{2T}} dy_1 \cdots dy_m \right)^{1/2}$$

and hence

$$\sup_{x \in K} II < \infty.$$

Thus, since factor I_n converges to 0 by assumption, we can approximate $\bar{u}$ uniformly in $x \in \mathbb{R}$ on compact sets by smooth pay-off functions. So $u \in C^1(\mathbb{R})$ and $u' = \bar{u}$. □

Next, we consider Asian options with pay-off given by (6.32). If $T_1 > 0$ we are able to give the analogous result to Theorem 6.23 by approximating the Asian pay-off with lookback pay-offs:

Corollary 6.24. *Let X^x be the strong solution to SDE* (6.5) *and $\Phi : \mathbb{R} \to \mathbb{R}$ a function in*

$L^{4p}_{\widetilde{w}}(\mathbb{R})$ where $\widetilde{w}$ is defined in (6.45) further below and where $p > 1$ is the conjugate of $1+\varepsilon$ for $\varepsilon > 0$ in Lemma 6.29. Then for any T_1, $T_2 \in (0, T]$ with $T_1 < T_2$, the price

$$u(x) = E\left[\Phi\left(\int_{T_1}^{T_2} X_u^x du\right)\right]$$

of the associated Asian option is continuously differentiable in $x \in \mathbb{R}$, and its derivative, i.e. the Delta, takes the form

$$u'(x) = E\left[\Phi\left(\int_{T_1}^{T_2} X_s^x ds\right)\int_0^{T_1} a(s)\frac{\partial}{\partial x}X_s^x dB_s\right] \tag{6.41}$$

for any bounded measurable function $a : \mathbb{R} \to \mathbb{R}$ such that

$$\int_0^{T_1} a(s)ds = 1. \tag{6.42}$$

Proof. Assume first that $\Phi \in C^1(\mathbb{R})$, and consider a series of partitions of $[T_1, T_2]$ with vanishing mesh, i.e. let $\{T_1 = t_0^m < t_1^m < \ldots < t_m^m = T_2\}_{m=1}^{\infty}$ with $\lim_{m\to\infty} \sup_{i=1,\ldots,m}(t_i^m - t_{i-1}^m) = 0$. Then we may write the integral using Riemann sums as follows

$$\int_{T_1}^{T_2} X_s^x ds = \lim_{m\to\infty} \sum_{i=1,\ldots,m} X_{t_i^m}^x (t_i^m - t_{i-1}^m).$$

Then

$$\Phi\left(\int_{T_1}^{T_2} X_s^x ds\right) = \lim_{m\to\infty} \Phi\left(\sum_{i=1,\ldots,m} X_{t_i^m}^x (t_i^m - t_{i-1}^m)\right) =: \lim_{m\to\infty} \hat{\Phi}_m(X_{t_1^m}^x, \ldots, X_{t_m^m}^x).$$

By Theorem 6.23 we have

$$u'(x) = \lim_{m\to\infty} E\left[\hat{\Phi}_m(X_{t_1^m}^x, \ldots, X_{t_m^m}^x)\int_0^T a_m(s)\frac{\partial}{\partial x}X_s^x dB_s\right]$$

where a_m is a bounded measurable function such that $\int_0^{t_i^m} a_m(s)ds = 1$ for each $i = 1, \ldots, m$. Then

$$\begin{aligned} u'(x) &= \lim_{m\to\infty} E\left[\hat{\Phi}_m(X_{t_1^m}^x, \ldots, X_{t_m^m}^x)\int_0^T a_m(s)\frac{\partial}{\partial x}X_s^x dB_s\right] \\ &= E\left[\Phi\left(\int_{T_1}^{T_2} X_s^x ds\right)\int_0^{T_1} a(s)\frac{\partial}{\partial x}X_s^x dB_s\right], \end{aligned}$$

where a is a function such that $\int_0^{T_1} a(s)ds = 1$.

For a general pay-off Φ, we approximate Φ in $L^{4p}_w(\mathbb{R})$ by a sequence of functions $\{\Phi_n\}_{n\geq 0} \subset C_0^1(\mathbb{R})$ and define $u(x) := E[\Phi(\int_{T_1}^{T_2} X_s^x ds)]$ and $\bar{u}(x) := E\left[\Phi\left(\int_{T_1}^{T_2} X_s^x ds\right)\int_0^{T_1} a(s)\frac{\partial}{\partial x}X_s^x dB_s\right]$. Consider $u_n(x) = E[\Phi_n(\int_{T_1}^{T_2} X_s^x ds)]$. Finally, similarly as in Theorem 6.23 one has $u_n(x) \to$

$u(x)$ as $n \to \infty$ for all $x \in \mathbb{R}$ and

$$|u_n'(x) - \bar{u}(x)| \lesssim E\left[\left|\Phi_n\left(\int_{T_1}^{T_2} B_s^x ds\right) - \Phi\left(\int_{T_1}^{T_2} B_s^x ds\right)\right|^{2p}\right]^{1/p},$$

which goes to zero uniformly in $x \in K$ on compact sets $K \subset \mathbb{R}$ as $n \to \infty$ by using the fact that $\int_{T_1}^{T_2} B_s^x ds$ has a Gaussian distribution with mean $x(T_2 - T_1)$ and variance $\frac{T_2^3 - T_1^3}{3} - (T_2 - T_1)T_1^2$ which explains the weight $\widetilde{w}$. $\square$

Remark 6.25. *From the proof of Corollary 6.24 it follows that the Delta* (6.41) *of an Asian option can be approximated by the Delta*

$$E\left[\Phi\left(\sum_{i=1}^{m} X_{t_i}^x (t_i - t_{i-1})\right) \int_0^{T_2} a(s) \frac{\partial}{\partial x} X_s^x dB_s\right] \tag{6.43}$$

of a lookback option for a fine enough partition $T_1 = t_0 < t_1 < \cdots < t_m = T_2$, where $\int_0^{t_i} a(s)ds = 1$ for each $i = 1, \ldots, m$.. From a numerical point of view, this might make a difference since the function a in (6.43) *can be chosen to have support on the full segment $[0, T_2]$, while in* (6.41) *the function a can only have support on $[0, T_1]$.*

If the averaging period of the Asian option starts today, i.e. $T_1 = 0$, then formula (6.41) does not hold anymore. Instead, one can derive alternative closed-form expressions for the Asian delta for smooth diffusion coefficients, see e.g. [50] and [18], which potentially can be generalised to irregular drift coefficients. However, except for linear coefficients (Black & Scholes model), these expressions involve stochastic integrals in the Skorokhod sense which are, in general, hard to simulate. Instead, we here propose to enlarge the state space by one dimension and to consider a perturbed Asian pay-off. In that case we are able to derive a probabilistic representation for the corresponding Delta that only includes Itô integrals. More precisely, we consider the (strong) solution to the perturbed two-dimensional SDE

$$dX_t^x = b(t, X_t^x)dt + dB_t,\ X_0^x = x \in \mathbb{R},$$

$$dY_t^{\epsilon,x,y} = X_t^x dt + \epsilon dW_t\ , Y_0^{\epsilon,x,y} = y \in \mathbb{R},\ 0 \leq t \leq T, \tag{6.44}$$

for $\epsilon > 0$, where W is a one-dimensional Brownian motion independent of B. The idea is now to consider the perturbed Asian pay-off with averaging period $[0, T_2]$, $T_2 \in (0, T]$ as a European pay-off on $Y_{T_2}^{\epsilon,x,y}$:

$$\Phi\left(\int_0^{T_2} X_s^x ds\right) \sim \Phi(Y_{T_2}^{\epsilon,x,0}) = \Phi\left(\int_0^{T_2} X_s^x ds + \epsilon W_{T_2}\right).$$

We then get the following result, where we now consider the slightly differently weighted pay-off function space

$$L_{\tilde{w}}^q(\mathbb{R}) := \left\{f : \mathbb{R} \to \mathbb{R} \text{ measurable: } \int_{\mathbb{R}} |f(x)|^q \tilde{w}(x)dx < \infty\right\}$$

for the weight function $\tilde{w}$ defined by

$$\tilde{w}(x) = \exp\left(-\frac{|x|^2}{2T_2\left(T_2^2/3+1\right)}\right), \quad x \in \mathbb{R}. \tag{6.45}$$

Theorem 6.26. *Let $Y_\cdot^{\epsilon,x,y}$ be the second component of the strong solution to (6.44) and $\Phi \in L_{\tilde{w}}^{4p}(\mathbb{R})$, where $p > 1$ is the conjugate of $1+\varepsilon$ for $\varepsilon > 0$ in Lemma 6.29. For a given maturity time $T_2 \in (0, T]$ and $0 < \epsilon \leq 1$, the price*

$$u_\epsilon(x) := E[\Phi(Y_{T_2}^{\epsilon,x,0})]$$

of the associated perturbed Asian option is continuously differentiable in $x \in \mathbb{R}$, and its derivative, i.e. the Delta, takes the form

$$u_\epsilon'(x) = E\left[\Phi(Y_{T_2}^{\epsilon,x,0})\left(\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s + \epsilon^{-1}\int_0^T a(s)\int_0^s \frac{\partial}{\partial x}X_u^x du\, dW_s\right)\right], \tag{6.46}$$

where $a : [0,T] \longrightarrow \mathbb{R}$ is a bounded measurable function such that $\int_0^T a(s)ds = 1$.

Proof. The proof is a straight forward generalization of the proof of Theorem 6.23 to the (particularly simple) two-dimensional extension (6.44) of the underlying SDE. Therefore, we here only give the main steps.

First observe that the strong solution $(X_t^x, Y_t^{\epsilon,x,y})$ is clearly differentiable in y, and by Theorem 6.17 also (weakly) differentiable in x, and we get

$$D_{x,y}\begin{pmatrix} X_t^x \\ Y_t^{\epsilon,x,y}\end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x}X_t^x & 0 \\ \int_0^t \frac{\partial}{\partial x}X_u^x du & 1\end{pmatrix},$$

for all $t \in [0,T]$, where $D_{x,y}$ denotes the derivative.

Assume first $\Phi \in C_0^\infty(\mathbb{R})$. Then it follows from Lemma 6.22 that $E[\Phi(Y_{T_2}^{\epsilon,x,y})]$ is continuously differentiable in (x,y) with

$$D_{x,y}E[\Phi(Y_{T_2}^{\epsilon,x,y})] = E\left[\begin{pmatrix} 0 \\ \Phi'(Y_{T_2}^{\epsilon,x,y})\end{pmatrix}^* D_{x,y}\begin{pmatrix} X_{T_2}^x \\ Y_{T_2}^{\epsilon,x,y}\end{pmatrix}\right],$$

where $*$ indicates the transposition of a matrix.

On the other hand, if we denote by D the Malliavin derivative in the direction of (B,W), it follows by means of the estimate in (6.78) and Corollary 6.19 that $Y_{T_2}^{\epsilon,x,y}$ is Malliavin differentiable and that for $0 \leq s \leq T$

$$D_s\begin{pmatrix} X_{T_2}^x \\ Y_{T_2}^{\epsilon,x,y}\end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \epsilon\end{pmatrix}^{-1} D_{x,y}\begin{pmatrix} X_s^x \\ Y_s^{\epsilon,x,y}\end{pmatrix} = D_{x,y}\begin{pmatrix} X_{T_2}^x \\ Y_{T_2}^{\epsilon,x,y}\end{pmatrix} \tag{6.47}$$

$dx \otimes ds \otimes P-$a.e. Then, using (6.47), the chain rule from Lemma 6.4 and the duality relation for the Malliavin derivative, we see that

$$D_{x,y}E[\Phi(Y_{T_2}^{\epsilon,x,y})] = E\left[\begin{pmatrix} 0 \\ \Phi'(Y_{T_2}^{\epsilon,x,y})\end{pmatrix}^* \int_0^T a(s) D_s \begin{pmatrix} X_{T_2}^x \\ Y_{T_2}^{\epsilon,x,y}\end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \epsilon\end{pmatrix}^{-1} D_{x,y}\begin{pmatrix} X_s^x \\ Y_s^{\epsilon,x,y}\end{pmatrix} ds\right]$$
$$= E\left[\Phi(Y_{T_2}^{\epsilon,x,y}) \int_0^T a(s) \left(\begin{pmatrix} 1 & 0 \\ 0 & \epsilon^{-1}\end{pmatrix} D_{x,y}\begin{pmatrix} X_s^x \\ Y_s^{\epsilon,x,y}\end{pmatrix}\right)^* d\begin{pmatrix} B_s \\ W_s\end{pmatrix}\right].$$

Thus

$$\frac{\partial}{\partial x}E[\Phi(Y_{T_2}^{\epsilon,x,y})] = E\left[\Phi(Y_{T_2}^{\epsilon,x,y})\left(\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s + \epsilon^{-1}\int_0^T a(s)\int_0^s \frac{\partial}{\partial x}X_u^x du\, dW_s\right)\right]$$

for all $x, y, \epsilon > 0$.

For general $\Phi \in L_{\tilde{w}}^{4p}(\mathbb{R})$ one pursues an approximation argument analogously to the one in the proof of Theorem 6.23, where we now use the Gaussian distribution of $\int_0^{T_2} B_s^x ds + \epsilon W_{T_2}$ with mean xT_2 and variance $T_2^3/3 + \epsilon^2 T_2$, which explains the weight (6.45) for $0 < \epsilon \leq 1$. □

Finally, we address the question whether both (6.41) for $T_1 \to 0$ as well as (6.46) for $\epsilon \to 0$ are indeed approximations for the Delta of the Asian option with averaging period starting in 0. We here give an affirmative answer for a class of pay-off functions Φ in spaces of the type

$$W_{\widetilde{w}}^{1,q}(\mathbb{R}) := \left\{ f \in W_{loc}^{1,q}(\mathbb{R});\ \int_{\mathbb{R}} |f(x)|^q \widetilde{w}(x)dx + \int_{\mathbb{R}} |f'(x)|^q \widetilde{w}(x)dx < \infty \right\}$$

for some $q > 1$, where f' denotes the weak derivative of f and the weight function $\widetilde{w}$ is defined in (6.45). See [69] for more information on weighted Sobolev spaces.

Theorem 6.27. *Let X^x be the strong solution to SDE* (6.5) *and $\Phi \in W_{\tilde{w}}^{1,4p}(\mathbb{R})$, where $p > 1$ is the conjugate of $1+\varepsilon$ for $\varepsilon > 0$ in Lemma 6.29. Further, require that the points of discontinuity of the distributional derivative Φ' are contained in a Lebesgue null set and that the following conditions are satisfied*

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \sup_{\epsilon>0} |\Phi(y) - \Phi(y-\epsilon z)|^{2p}\, \widetilde{w}(y) P_T(z) dy dz < \infty \tag{6.48}$$

and

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \sup_{\epsilon>0} |\Phi'(y) - \Phi'(y-\epsilon z)|^{4p}\, \widetilde{w}(y) P_T(z) dy dz < \infty, \tag{6.49}$$

where $P_t(z) = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{1}{2t}z^2)$, $t > 0$, $z \in \mathbb{R}$ is the Gaussian kernel. Then

$$u(x) := E\left[\Phi\left(\int_0^{T_2} X_s^x ds\right)\right]$$

is continuously differentiable in $x \in \mathbb{R}$, and

$$u'(x) = \lim_{\epsilon\to 0} E\left[\Phi\left(Y_{T_2}^{\epsilon,x,0}\right)\left(\int_0^T a(s)\frac{\partial}{\partial x}X_s^x dB_s + \epsilon^{-1}\int_0^T a(s)\int_0^s \frac{\partial}{\partial x}X_u^x du\, dW_s\right)\right], \tag{6.50}$$

as well as

$$u'(x) = \lim_{T_1 \to 0} E\left[\Phi\left(\int_{T_1}^{T_2} X_u^x du\right) \int_0^{T_1} a(u) \frac{\partial}{\partial x} X_u^x dB_u\right]. \tag{6.51}$$

Proof. By Theorem 6.26 we have that $u_\epsilon \in C^1(\mathbb{R})$ for all $\epsilon > 0$. Hence,

$$\frac{\partial}{\partial x} E[\Phi(Y_{T_2}^{\epsilon,x,0})] = E\left[\Phi'(Y_{T_2}^{\epsilon,x,0}) \frac{\partial}{\partial x} Y_{T_2}^{\epsilon,x,0}\right]$$

for all $\epsilon > 0$, dx-a.e. Let $J \subset \mathbb{R}$ be a compact set. Then, using the same line of reasoning just as in the proof of Theorem 6.23, using Cauchy-Schwarz inequality, Girsanov's theorem, and Lemma 6.29 we find the estimates

$$\sup_{x \in J} \left| \frac{\partial}{\partial x} E[\Phi(Y_{T_2}^{\epsilon,x,0})] - \frac{\partial}{\partial x} E[\Phi(Y_{T_2}^{0,x,0})] \right| \le C \left(E\left[\int_{\mathbb{R}} \left| \Phi'(y) - \Phi'(y - \epsilon W_{T_2}) \right|^{4p} \tilde{w}(y) dy \right] \right)^{\frac{1}{4p}}$$

and

$$\sup_{x \in J} \left| E[\Phi(Y_{T_2}^{\epsilon,x,0})] - E[\Phi(Y_{T_2}^{0,x,0})] \right| \le K \left(E\left[\int_{\mathbb{R}} \left| \Phi(y) - \Phi(y - \epsilon W_{T_2}) \right|^{2p} \tilde{w}(y) dy \right] \right)^{\frac{1}{2p}}$$

for constants C, K depending only on T_2, J, p (and not on ϵ).

Finally, using dominated convergence in connection with (6.48) and (6.49), the proof follows.

To prove (6.51) define $u_{T_1}(x) := E[\Phi(\int_{T_1}^{T_2} X_u^x du)]$. Since $\Phi \in L_{\tilde{w}}^{4p}(\mathbb{R})$, we have by Corollary 6.24 that $u_{T_1} \in C^1(\mathbb{R})$ for every $T_1 > 0$. Moreover, since $\Phi \in W_{\tilde{w}}^{1,4p}(\mathbb{R})$ we have

$$u'_{T_1}(x) = E\left[\Phi'\left(\int_{T_1}^{T_2} X_u^x du\right) \int_{T_1}^{T_2} \frac{\partial}{\partial x} X_u^x du\right].$$

Consequently, for every compact $J \subset \mathbb{R}$ we have

$$\begin{aligned}
&\sup_{x \in J} \left| \frac{\partial}{\partial x} E\left[\Phi\left(\int_{T_1}^{T_2} X_u^x du\right)\right] - \frac{\partial}{\partial x} E\left[\Phi\left(\int_0^{T_2} X_u^x du\right)\right] \right| \\
&\le \sup_{x \in J} \left| E\left[\left(\Phi'\left(\int_{T_1}^{T_2} X_u^x du\right) - \Phi'\left(\int_0^{T_2} X_u^x du\right)\right) \int_{T_1}^{T_2} \frac{\partial}{\partial x} X_u^x du\right] \right| \\
&\quad + \sup_{x \in J} \left| E\left[\Phi'\left(\int_0^{T_2} X_u^x du\right) \int_0^{T_1} \frac{\partial}{\partial x} X_u^x du\right] \right| \\
&=: A_1 + A_2
\end{aligned}$$

where A_1 and A_2 denote the respective summands. It is clear that A_2 goes to 0 uniformly in x on J as $T_1 \to 0$. To show the corresponding convergence for A_1, similar computations as in the beginning of the proof, using Cauchy-Schwarz inequality, Girsanov's theorem, Lemma 6.29, and that $\Phi' \in L^{4p}_{\tilde{w}}(\mathbb{R})$, give for some constant $C_\varepsilon > 0$

$$A_1 = C_\varepsilon \sup_{x \in J} E\left[\left|\Phi'\left(\int_0^{T_1} B_u^x du\right)\right|^{4p}\right]^{1/(4p)} \leq C_\varepsilon \|\Phi'\|_{L^{4p}_{\tilde{w}}(\mathbb{R})} \int_{\mathbb{R}} e^{-\frac{z^2}{2T_1(T_1^2/3+1)}} dz \xrightarrow{T_1 \to 0} 0\,.$$

Hence (6.51) follows. □

Example 6.28. *We conclude this section by verifying the conditions in Theorem 6.27 for a payoff function that is used in the next section. Consider the function $\Phi : \mathbb{R} \longrightarrow [0, \infty)$ given by*

$$\Phi(y) = \exp(-y)(C\exp(y) - K)_+,$$

where $C,\ K > 0$ are constants and $(x)_+ := \max(x, 0)$ for $x \in \mathbb{R}$. We immediately see that $\Phi \in W^{1,4p}_{loc}(\mathbb{R}) \cap L^{4p}_{\tilde{w}}(\mathbb{R})$ and that

$$\Phi^{'}(y) = -\exp(-y)(C\exp(y) - K)_+ + C\mathbf{1}_{[\log(K/C),\infty)}(y)\ dx - a.e.$$

On the other hand we have that

$$\begin{aligned}\sup_{\epsilon>0}\left|\Phi^{'}(y) - \Phi^{'}(y-\epsilon z)\right|^{4p} &\leq M(\left|\Phi^{'}(y)\right|^{4p} + \sup_{\epsilon>0}\left|\Phi^{'}(y-\epsilon z)\right|^{4p} \\ &\leq M((2C + K\exp(|y|))^{4p} + (2C + K\exp(|y| + |z|))^{4p}).\end{aligned}$$

So condition (6.49) is fulfilled. In the same way one verifies condition (6.48). Hence Φ satisfies the assumptions of the previous theorem.

6.5 Examples and Simulations

We complete this paper by applying the results from Section 6.4 to the computation of the Deltas in the regime-switching examples mentioned in the Introduction. More complex examples of state-dependent drift coefficients (see e.g. [35]) can be treated following the same principles. To implement the methodology, we first employ Remark 6.20 and observe that all drift coefficients from the regime switching examples in the Introduction can be written in the form $b(t,x) = \tilde{b}(x) + \hat{b}(x)$ as in (6.6) such that identity (6.20) holds for $\hat{b}(x)$ and identity (6.22) holds for $\tilde{b}(x)$. We thus get the following rewriting of the first variation process (6.27):

$$\frac{\partial}{\partial x} X_t^x = \exp\left\{2\tilde{\beta}(X_t^x) - 2\tilde{\beta}(x) - 2\int_0^t \tilde{b}(X_s^x)\, dX_s^x + \int_0^t \hat{b}'(X_u^x) du\right\}\,, \tag{6.52}$$

where $\tilde{\beta}(\cdot) := \tilde{b}(0) + \int_0^{\cdot} \tilde{b}(y)\, dy$ is a primitive of $\tilde{b}(\cdot)$. This form is convenient for simulation purposes.

6.5.1 Black & Scholes model with regime-switching dividend yield

Consider an extended Black & Scholes model where the stock pays a dividend yield that switches to a higher level when the stock value passes a certain threshold $\overline{R} \in \mathbb{R}_+$. That is, under the risk-neutral measure the stock price S is given by the SDE

$$S_t^{s_0} = s_0 + \int_0^t \overline{b}(S_u^{s_0}) S_u^{s_0} du + \int_0^t \sigma S_u^{s_0} dB_u \,,$$

where $\sigma > 0$ is constant and the drift coefficient $\overline{b} : \mathbb{R} \to \mathbb{R}$ is given by

$$\overline{b}(x) := -\overline{\lambda}_1 \mathbf{1}_{(-\infty,\overline{R})}(x) - \overline{\lambda}_2 \mathbf{1}_{[\overline{R},\infty)}(x),$$

for dividend yields $\overline{\lambda}_1, \overline{\lambda}_2 \in \mathbb{R}_+$. We are interested in computing the Delta of a European option written on the stock with given pay-off function $\overline{\Phi} : \mathbb{R} \to \mathbb{R}$ and maturity T:

$$\frac{\partial}{\partial s_0} E[\overline{\Phi}(S_T^{s_0})] \,.$$

In order to fit the computation of the Delta in our framework, we rewrite the stock price with the help of Itô's formula as

$$S_T^{s_0} = e^{\sigma X_T^{\ln(s_0)/\sigma}} \,,$$

where X_t^x is the solution of the SDE

$$X_t^x = x + \int_0^t b(X_u^x) du + B_t \,, \tag{6.53}$$

with

$$b(x) := -\lambda_1 \mathbf{1}_{(-\infty,R)}(x) - \lambda_2 \mathbf{1}_{[R,\infty)}(x) - \frac{\sigma}{2},$$

and $\lambda_1 := \frac{\overline{\lambda}_1}{\sigma}$, $\lambda_2 := \frac{\overline{\lambda}_2}{\sigma}$, $R := \frac{\ln(\overline{R})}{\sigma}$. We see that SDE (6.53) is in the required form (6.5) with $\tilde{b}(t,x) = -(\lambda_2 - \lambda_1)\mathbf{1}_{[R,\infty)}(x)$ and $\hat{b}(t,x) = -\lambda_1 - \frac{\sigma}{2}$. With $\Phi := \overline{\Phi} \circ \exp \circ \sigma \cdot$ we thus get by the chain rule

$$\frac{\partial}{\partial s_0} E[\overline{\Phi}(S_T^{s_0})] = \frac{\partial}{\partial s_0} E[\Phi(X_T^{\ln(s_0)/\sigma})] = \frac{1}{s_0 \sigma} \cdot \frac{\partial}{\partial x} E[\Phi(X_T^x)] \,|_{x=\frac{\ln(s_0)}{\sigma}} \,.$$

If $\Phi \in L_w^{4p}(\mathbb{R})$ we know by Theorem 6.23 that the Delta exists, and we can compute $\frac{\partial}{\partial x} E[\Phi(X_T^x)]$ by (6.38) to obtain

$$\frac{\partial}{\partial s_0} E[\overline{\Phi}(S_T^{s_0})] = E\left[\overline{\Phi}(S_T^{s_0}) \int_0^T \frac{a(s)}{s_0 \sigma} \frac{\partial}{\partial x} X_s^{\ln(s_0)/\sigma} dB_s \right] \tag{6.54}$$

for any bounded measurable function $a : \mathbb{R} \to \mathbb{R}$ such that $\int_0^T a(s) ds = 1$, and where $\frac{\partial}{\partial x} X_s^x$ is given by (6.52) with $\hat{b}' = 0$ and

$$\tilde{\beta}(x) := \int_0^x \tilde{b}(y)\, dy = -(\lambda_2 - \lambda_1)(x - R)\mathbf{1}_{[R,\infty)}(x).$$

We now consider the Delta for a call option, i.e. $\overline{\Phi}(x) := (x-K)_+$, and for a digital option, i.e. $\overline{\Phi}(x) := \mathbf{1}_{\{x\geq K\}}$, for some strike price $K>0$. It is easily seen that in both cases $\Phi \in L^{4p}_w(\mathbb{R})$. To compute (6.54) by Monte Carlo, X^x is approximated by an Euler scheme (see [109], Theorem 3.1 on the Euler scheme approximation for coefficients b which are non-Lipschitz due to a set of discontinuity points with Lebesgue measure zero). As in [50] we compare the performance of (6.54) to the approximation of the Delta by a finite difference scheme combined with Monte Carlo:

$$\frac{\partial}{\partial s_0}E[\overline{\Phi}(S_T^{s_0})] \sim \frac{E[\overline{\Phi}(S_T^{s_0+\epsilon})] - E[\overline{\Phi}(S_T^{s_0-\epsilon})]}{2\epsilon}, \tag{6.55}$$

for ϵ sufficiently small. We set the parameters $T=1$, $s_0=100$, $\bar{\lambda}_1=0.05$, $\bar{\lambda}_2=0.15$, $\bar{R}=108$, $\sigma=0.1$ and $K=94$. Our findings are analogue to the ones in [50]: for the continuous call option pay-off function the approximation (6.55) seems to be more efficient (see Figure 6.1), whereas for the discontinuous pay-off function of a digital option, the approximation (6.54) via the Malliavin weight exhibits considerably better convergence (see Figure 6.2).

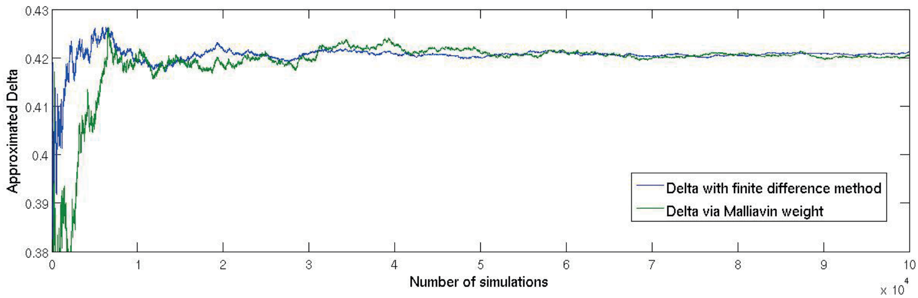

Figure 6.1: Delta of a European Call Option Black & Scholes model with regime-switching dividend yield.

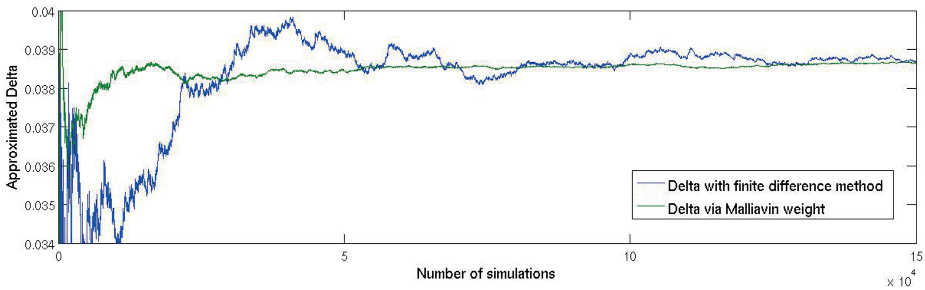

Figure 6.2: Delta of a European Digital Option under the Black & Scholes model with regime-switching dividend yield.

6.5.2 Electricity spot price model with regime-switching mean-reversion rate

Typically, electricity spot prices exhibit a mean-reverting behaviour with at least two different regimes of mean-reversion: a spike regime with very strong mean-reversion on exceptionally high price levels and a base regime with moderate mean-reversion on regular price levels. These features can be captured by modelling the electricity spot price S (under a risk-neutral pricing measure) by an extended Ornstein-Uhlenbeck process with regime-switching mean-reversion rate:

$$S_t^{s_0} = s_0 + \int_0^t \overline{b}(S_u^{s_0})du + \sigma B_t\,, \tag{6.56}$$

where the drift coefficient is given by

$$\overline{b}(x) := -\overline{\lambda}_1 x \mathbf{1}_{(-\infty,\overline{R})}(x) - \overline{\lambda}_2 x \mathbf{1}_{[\overline{R},\infty)}(x) \tag{6.57}$$

for mean reversion rates $\overline{\lambda}_1, \overline{\lambda}_2 \in \mathbb{R}_+$, a given spike price threshold $\overline{R} \in \mathbb{R}$, and $\sigma > 0$. In order to guarantee positive prices, one could alternatively model the log-price by (6.56), or one could introduce another regime with high mean-reversion as soon as the price falls below zero (we recall that short periods of negative electricity prices have been observed).

Since electricity is a flow commodity, derivatives on spot electricity are written on the average price of the delivery of 1 kWh over a future period $[T_1, T_2]$, i.e. the underlying is of the type $\int_{T_1}^{T_2} S_t^{s_0}\,dt$ for $T_1 > 0$. The most liquidly traded electricity derivatives are futures and forwards. In that case the pay-off is linear and the computation of the Delta can be reduced to the computation of the Deltas of European type options:

$$\frac{\partial}{\partial s_0} E\left[\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} S_t^{s_0}\,dt\right] = \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} \frac{\partial}{\partial s_0} E[S_t^{s_0}]\,dt\,.$$

For derivatives with non-linear pay-off Φ, the Delta

$$\frac{\partial}{\partial s_0} E\left[\Phi\left(\int_{T_1}^{T_2} S_t^{s_0}\,dt\right)\right]$$

is of Asian type.

Again, in order to fit the computation of the Delta in our framework we rewrite the stock price with the help of Itô's formula as

$$S_t^{s_0} = \sigma X_t^{s_0/\sigma},$$

where X^x is the solution of the SDE

$$X_t^x = x + \int_0^t b(X_u^x)du + B_t, \tag{6.58}$$

with

$$b(x) := -\left(\lambda_1 \mathbf{1}_{(-\infty,R)}(x) + \lambda_2 \mathbf{1}_{[R,\infty)}(x)\right) x, \tag{6.59}$$

where $R = \overline{R}/\sigma$. We see that the SDE (6.58) is in the required form (6.5) with $\tilde{b}(x) =$

$-(\lambda_2 - \lambda_1) R\mathbf{1}_{[R,\infty)}(x)$ and $\hat{b}(x) = b(x) - \tilde{b}(x)$. As in the previous example, by the chain rule we get that

$$\frac{\partial}{\partial s_0} E\left[\Phi\left(\int_{T_1}^{T_2} S_t^{s_0} dt\right)\right] = \frac{\partial}{\partial s_0} E\left[\overline{\Phi}\left(\int_{T_1}^{T_2} X_t^{s_0/\sigma} dt\right)\right] = \frac{1}{\sigma}\frac{\partial}{\partial x} E\left[\overline{\Phi}\left(\int_{T_1}^{T_2} X_t^{x} dt\right)\right]\Bigg|_{x=\frac{s_0}{\sigma}} \tag{6.60}$$

with $\overline{\Phi} := \Phi \circ \cdot\sigma$. Note that in this example the first variation process $\frac{\partial}{\partial x} X_s^x$ is given by (6.52) with

$$\tilde{\beta}(x) := \int_0^x \tilde{b}(y)\, dy = -(\lambda_2 - \lambda_1) R(x - R)\mathbf{1}_{[R,\infty)}(x)$$

and

$$\int_0^t \hat{b}'(u, X_u^x) du = -\lambda_1 \int_0^t \mathbf{1}_{(-\infty,R)}(X_u^x) du - \lambda_2 \int_0^t \mathbf{1}_{[R,\infty)}(X_u^x) du.$$

We compare the performance of the formula for the Asian Delta in Corollary 6.24 with the approximation presented in Remark 6.25 and with a finite difference approximation analogous to (6.55) when Φ is a call option pay-off and a digital option pay-off, respectively. Obviously, in both cases the pay-off in terms of $X_\cdot^x$ fulfils the assumptions in Theorem 6.23. In the approximation presented in Remark 6.25 an optimal (in the sense that it minimises the variance of the Malliavin weight) choice for $a(s)$ could improve the convergence rate of the method. In the simulations we compared the following possible choices for $a(s)$:

$$a_1(s) := \begin{cases} \frac{1}{t_1} & \text{if } \ 0 \le s \le t_1 \\ 0 & \text{if } \ t_1 < s \le T_2 \end{cases}$$

$$a_2(s) := \begin{cases} \frac{1}{t_1} & \text{if } \ 0 \le s \le t_1 \\ k & \text{if } \ \left\lfloor \frac{s-T_1}{T_2-T_1} \cdot 2m \right\rfloor \equiv 0 \mod 2 \ \text{ and } t_1 < s \le T_2 \\ -k & \text{if } \ \left\lfloor \frac{s-T_1}{T_2-T_1} \cdot 2m \right\rfloor \equiv 1 \mod 2 \ \text{ and } t_1 < s \le T_2 \end{cases}$$

$$a_3(s) := \begin{cases} \frac{1}{t_1} & \text{if } \ 0 \le s \le t_1 \\ \left| \frac{s-T_1}{T_2-T_1} \cdot \frac{m}{2} - 1 - \left\lfloor \frac{s-T_1}{T_2-T_1} \cdot \frac{m}{2} - \frac{1}{2} \right\rfloor \right| - k & \text{if } \ t_1 < s \le T_2 \end{cases},$$

see Figure 6.3. However, the different choices of function a above did not produce relevant differences in the results. Note, that implementing the approximation from Remark 6.25 with function $a_1(s)$ is essentially the same as the implementing the Delta from Corollary 6.24. We thus only compare the Delta from Corollary 6.24 with a finite difference scheme for parameters: $T_1 = 0.4$, $T_2 = 1$, $s_0 = 100$, $\overline{\lambda}_1 = 0.2$, $\overline{\lambda}_2 = 0.4$, $\overline{R} = 101$, $\sigma = 5$ and $K = 87$. We remark that if T_1 approaches zero, the variance of the Malliavin weight increases, and thereby the Monte Carlo method becomes less effective. As for the European option in Subsection 6.5.1, also for these Asian type options the finite difference method seems to be more efficient for the continuous call option pay-off, see Figure 6.4, whereas for the digital option pay-off, the approximation through the Malliavin weight provides better convergence, see Figure 6.5.

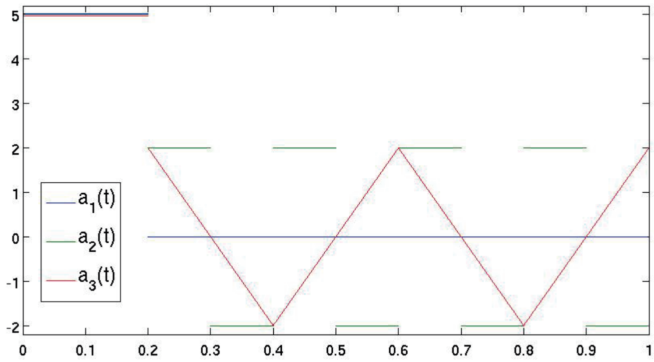

Figure 6.3: Three versions of the functions for $a(s)$ from Remark 6.25.

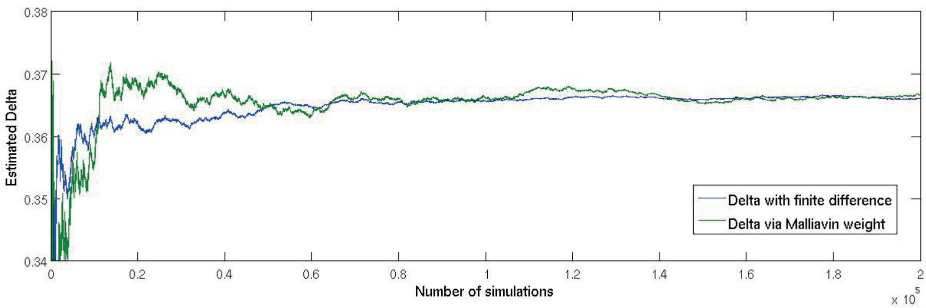

Figure 6.4: Delta of an Asian Call Option under the Electricity spot price model with regime-switching mean-reversion rate.

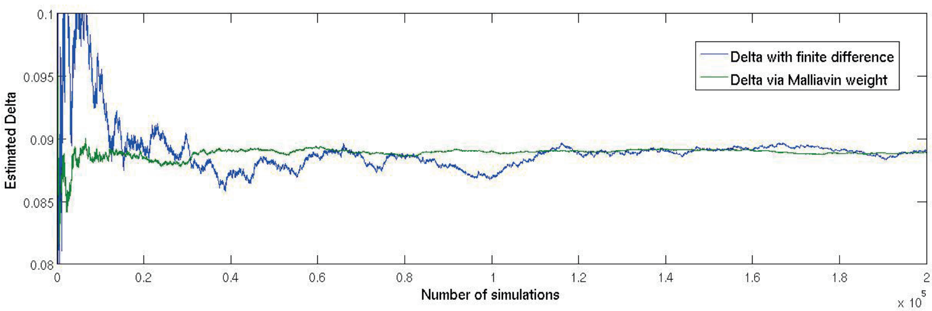

Figure 6.5: Delta of an Asian Digital Option under the Electricity spot price model with regime-switching mean-reversion rate.

6.5.3 Generalised Black & Scholes model with regime-switching short rate

Consider a generalised Black & Scholes model where under the risk-neutral measure the stock price $S^{s_0}_{\cdot}$ is given by

$$S^{s_0}_t = s_0 + \int_0^t r^{r_0}_u S^{s_0}_u du + \int_0^t \sigma S^{s_0}_u dB_u \,, \tag{6.61}$$

and the stochastic short rate $r^{r_0}_{\cdot}$ is given by an extended Vašíček model where the mean-reversion level switches between a high interest rate regime and a low interest rate regime when the short rate passes a certain threshold $R \in \mathbb{R}$:

$$r^{r_0}_t = r_0 + \int_0^t b(r^{r_0}_u) du + B^*_t \,, \tag{6.62}$$

where $B^*_t = \rho \widetilde{B}_t + \sqrt{1-\rho^2} B_t$ and the drift coefficient is given by

$$b(x) := -\lambda (x - m_1 \mathbf{1}_{(-\infty,R)}(x) - m_2 \mathbf{1}_{[R,\infty)}(x)) \tag{6.63}$$

for a mean-reversion rate $\lambda \in \mathbb{R}_+$ and mean-reversion levels $m_1, m_2 \in \mathbb{R}$, and where $\widetilde{B}$ is a Brownian motion independent of B, i.e. we allow for a correlation coefficient $0 \leq \sqrt{1-\rho^2} < 1$ with the stock price. Note that we set the volatility coefficient in (6.62) equal to one for notational simplicity. We see that the drift of the SDE (6.62) is in the required form (6.5) with $\tilde{b}(x) = -\lambda\,(m_1 - m_2)\,\mathbf{1}_{[R,\infty)}(x)$ and $\hat{b}(x) = -\lambda\,(x - m_1)$. Further, we mention that the SDE (6.62) has a Malliavin differentiable unique strong solution with respect to the filtration $\mathcal{F}_t, 0 \leq t \leq T$, generated by the Brownian motions $\widetilde{B}_{\cdot}$ and $B_{\cdot}$. Moreover, there exists an Ω^* with probability mass 1 such that for all $\omega \in \Omega^*$ and $0 \leq t \leq T : (x \longmapsto r^x(t,\omega)) \in \cap_{p>0} W^{1,p}_{loc}(\mathbb{R})$. The proofs of these properties are essentially the same as in Section 6.3. For example, Girsanov's theorem in the previous proofs is applied to the Brownian motion $B^*_t := \rho \widetilde{B}_t + \sqrt{1-\rho^2} B_t, 0 \leq t \leq T$.

Now consider the price of a European option with pay-off function Φ written on the stock at maturity T:

$$E\left[e^{-\int_0^T r^{r_0}_s ds} \Phi\left(s_0 e^{\int_0^T r^{r_0}_s ds + \sigma B_T - \frac{1}{2}\sigma^2 T}\right)\right] .$$

In this example we are interested in computing the *generalised Rho*

$$\frac{\partial}{\partial r_0} E\left[e^{-\int_0^T r^{r_0}_s ds} \Phi\left(s_0 e^{\int_0^T r^{r_0}_s ds + \sigma B_T - \frac{1}{2}\sigma^2 T}\right)\right] , \tag{6.64}$$

that is, the sensitivity of the option with respect to the initial value r_0 of the short rate (i.e. a sensitivity with respect to movements of the short end of the yield curve). We see that (6.64) has the form of a Delta with respect to an Asian pay-off in the short rate $r^{r_0}_{\cdot}$ which, however, additionally depends on the factor B_T.

Although the extension of the results in Section 6.4 is straight forward to this simple two-dimensional setting, we can still remain in the one-dimensional setting from Section 6.4 by considering the Malliavin derivative $\widetilde{D}_s$ only with respect to Brownian motion $\widetilde{B}_{\cdot}$ and by applying relation (6.28) from Corollary 6.19 in the form

$$\frac{\partial}{\partial r_0} r_t^{r_0} = \frac{1}{\rho} \widetilde{D}_s r_t^{r_0} \frac{\partial}{\partial r_0} r_s^{r_0} \text{ for all } s \leq t\,. \tag{6.65}$$

We here intend to analyse the performance of the approximation (6.50) from Theorem 6.27 for an Asian Delta. Under the corresponding assumptions from Theorem 6.27 for the pay-off function

$$\bar{\Phi}\left(\int_0^T r_t^{r_0} dt, B_T\right) := \exp\left\{-\int_0^T r_t^{r_0} dt\right\} \Phi\left(s_0 \exp\left\{\int_0^T r_t^{r_0} dt + \sigma B_T - \frac{1}{2}\sigma^2 T\right\}\right),$$

and by following the argument in the proof of Theorem 6.27 we then obtain that the function

$$u(r_0) := E\left[\bar{\Phi}\left(\int_0^T r_t^{r_0} dt, B_T\right)\right]$$

is continuously differentiable in $r_0 \in \mathbb{R}$, and that

$$\begin{aligned} \frac{\partial}{\partial r_0} u(r_0) = \lim_{n\to\infty} E\Bigg[\bar{\Phi}\left(\int_0^T r_s^{r_0} ds + n^{-1} W_T, B_T\right) \\ \left(\int_0^T \frac{a(s)}{\rho} \frac{\partial}{\partial r_0} r_s^{r_0} d\widetilde{B}_s + n \int_0^T a(s) \left(\int_0^s \frac{\partial}{\partial r_0} r_u^{r_0} du\right) dW_s\right)\Bigg] \end{aligned} \tag{6.66}$$

where $a : \mathbb{R} \to \mathbb{R}$ is as in Theorem 6.27. Note that in this example the first variation process $\frac{\partial}{\partial r_0} r_s^{r_0}$ is given by (6.52) with

$$\tilde{\beta}(x) := \int_0^x \tilde{b}(y)\, dy = -\lambda\,(m_1 - m_2)\,(x - R)\,\mathbf{1}_{[R,\infty)}(x)$$

and

$$\int_0^t \hat{b}'(u, X_u^x) du = -\lambda t.$$

We compare the performance of the approximation of the generalised Rho $\frac{\partial}{\partial r_0} u$ presented in (6.66) with a finite difference approximation analogous to (6.55) when Φ is a call option pay-off, see Figure 6.6. The parameters are $T = 1$, $s = 2$, $\sigma = 0.1$, $\lambda = 0.3$, $m_1 = 0.5$, $m_2 = 1.2$, $R = 1.4$ and $K = \exp(0.4)$ and we choose $a(s) = 1/T$. Note that for a call option pay-off Φ we know from Example 6.28 that the assumptions in Theorem 6.27 are fulfilled. Further, we also compute the Delta of a digital pay-off, see Figure 6.7, even though the conditions of Theorem 6.27 are not satisfied. Our conjecture is that the result of Theorem 6.27 also holds for discontinuous pay-offs, and the simulation reinforces that. As n from Theorem 6.27 increases $\bar{\Phi}\left(\int_0^T r^{r_0}(s)ds + n^{-1}W_T, B_T\right)$ becomes a better approximation of $\bar{\Phi}\left(\int_0^T r^{r_0}(s)ds, B_T\right)$ but at the same time the variance of the Malliavin weight increases, thus, the convergence of the Monte Carlo simulation becomes slower. The experience of several simulations is that $n \sim 20$ gives the best balance between these two opposite impacts. However, we can see that in both cases the finite difference method seems considerably more efficient.

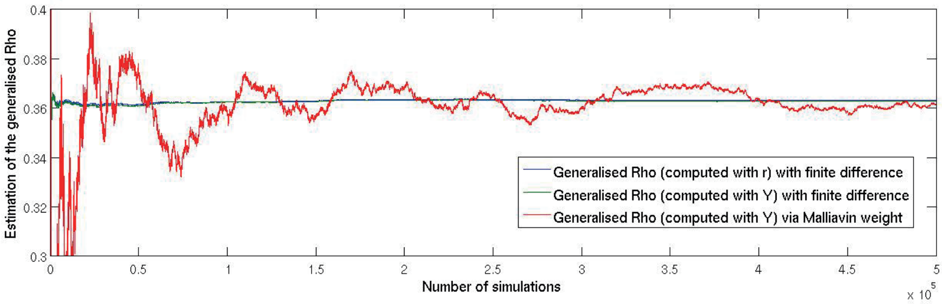

Figure 6.6: Approximation: Generalised Rho of a European Call Option under the Generalised Black & Scholes model with regime-switching short rate.

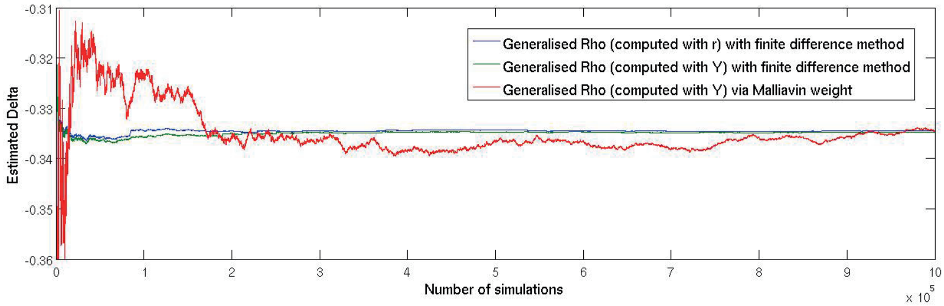

Figure 6.7: Approximation: Generalised Rho of a European Digital Option under the Generalised Black & Scholes model with regime-switching short rate.

Appendix

6.A Proofs of results in Section 6.3

In this appendix we recollect the proofs of the results in Section 6.3.

6.A.1 Some auxiliary results

We start by giving some auxiliary technical lemmata which provide relevant estimates that will be progressively used throughout some proofs in the sequel.

Lemma 6.29. *Let $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a function of at most linear growth, i.e. $|b(t,x)| \leq C(1+|x|)$ for some $C > 0$, all $x \in \mathbb{R}$ and $t \in [0,T]$. Then for any compact subset $K \subset \mathbb{R}$ there exists an $\varepsilon > 0$ such that*

$$\sup_{x \in K} E\left[\mathcal{E}\left(\int_0^T b(u, B_u^x) dB_u\right)^{1+\varepsilon}\right] < \infty \tag{6.67}$$

where $B_t^x := x + B_t$.

Proof. Indeed, write

$$\begin{aligned}
E&\left[\mathcal{E}\left(\int_0^T b(u, B_u^x) dB_u\right)^{1+\varepsilon}\right] = \\
&= E\left[\exp\left\{\int_0^T (1+\varepsilon) b(u, B_u^x) dB_u - \frac{1}{2}\int_0^T (1+\varepsilon) b^2(u, B_u^x) du\right\}\right] \\
&= E\left[\exp\left\{\int_0^T (1+\varepsilon) b(u, B_u^x) dB_u - \frac{1}{2}\int_0^T (1+\varepsilon)^2 b^2(u, B_u^x) du \right.\right. \\
&\quad \left.\left. + \frac{1}{2}\int_0^T \varepsilon(1+\varepsilon) b^2(u, B_u^x) du\right\}\right] \\
&= E\left[\exp\left\{\frac{1}{2}\int_0^T \varepsilon(1+\varepsilon) b^2(u, X_u^{\varepsilon,x}) du\right\}\right]
\end{aligned}$$

where in the last step $X^{\varepsilon,x}$ denotes a weak solution of the SDE

$$\begin{cases} dX_t^{\varepsilon,x} = (1+\varepsilon)b(t,X_t^{\varepsilon,x})dt + dB_t, \quad t\in[0,T] \\ X_0^{\varepsilon,x} = x, \end{cases}$$

which is obtained from Girsanov's theorem in the same way as in the first step of Subsection 6.A.2 in equation (6.74). Observe that, since b has at most linear growth, we have

$$|X_t^{\varepsilon,x}| \le |x| + C(1+\varepsilon)\int_0^t (1+|X_u^{\varepsilon,x}|)du + |B_t|$$

for every $t\in[0,T]$. Then Grönwall's inequality gives

$$|X_t^{\varepsilon,x}| \le (|x| + C(1+\varepsilon)T + |B_t|)\, e^{C(1+\varepsilon)T}, \tag{6.68}$$

and due to the sublinearity of b and the estimate (6.68) we can find a constant $C_{\varepsilon,T}$ depending only on ε, T such that $\lim_{\varepsilon\searrow 0} C_{\varepsilon,T} < \infty$ and

$$|b(u,X_u^{\varepsilon,x})| \le C_{\varepsilon,T}\left(1+|x|+|B_t|\right).$$

As a result,

$$\begin{aligned} E\Big[\exp\Big\{\varepsilon(1+\varepsilon)\int_0^T b^2(u,X_u^{\varepsilon,x})du\Big\}\Big] &\le E\left[\exp\left\{\varepsilon(1+\varepsilon)C_{\varepsilon,T}^2\int_0^T (1+|x|+|B_u|)^2\,du\right\}\right] \\ &\le e^{\tilde{C}_{\varepsilon,T}T(1+|x|)^2}E\left[\exp\left\{2\tilde{C}_{\varepsilon,T}(1+|x|)\int_0^T |B_u|du + \tilde{C}_{\varepsilon,T}\int_0^T |B_u|^2du\right\}\right] \end{aligned}$$

where $\tilde{C}_{\varepsilon,T} := \varepsilon(1+\varepsilon)C_{\varepsilon,T}^2 > 0$ is a constant such that $\lim_{\varepsilon\searrow 0}\tilde{C}_{\varepsilon,T} = 0$. Clearly, from the above expression we can see that for every compact set $K\subset\mathbb{R}$ we can choose $\varepsilon>0$ small enough such that

$$\sup_{x\in K} E\Big[\exp\Big\{\varepsilon(1+\varepsilon)\int_0^T b^2(u,X_u^{\varepsilon,x})du\Big\}\Big] < \infty.$$

□

Remark 6.30. *From Lemma 6.29 it follows immediately that if the approximating functions b_n, $n\ge 1$ are as in (6.23) then for any compact subset $K\subset\mathbb{R}$, one can find an $\varepsilon>0$ such that*

$$\sup_{x\in K}\sup_{n\ge 0} E\left[\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)^{1+\varepsilon}\right] < \infty, \tag{6.69}$$

where we recall that $b_0 := b$.

Lemma 6.31. *Let $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ be a bounded measurable function. Then for every $t\in[0,T]$, $\lambda\in\mathbb{R}$ and compact subset $K\subset\mathbb{R}$ we have*

$$\sup_{x\in K} E\left[\exp\left\{\lambda\int_0^t\int_{\mathbb{R}} f(s,y)L^{B^x}(ds,dy)\right\}\right] < \infty \tag{6.70}$$

where $L^{B^x}(ds,dy)$ denotes integration with respect to local-time of the Brownian motion $B_t^x := B_t + x$ in both time and space, see Section 6.2 or [39] for more information on local-time integration.

Proof. By virtue of decomposition (6.19) from the Section 6.2 and Cauchy-Schwarz inequality twice we have

$$\begin{aligned}
E\left[\exp\left\{\lambda\int_0^t\int_{\mathbb{R}} f(s,y)L^{B^x}(ds,dy)\right\}\right] &\le E\left[\exp\left\{-2\lambda\int_0^t f(s,B_s^x)dB_s\right\}\right]^{1/2}\\
&\times E\left[\exp\left\{4\lambda\int_{T-t}^T f(T-s,B_{T-s}^x)dW_s\right\}\right]^{1/4}\\
&\times E\left[\exp\left\{-4\lambda\int_{T-t}^T f(T-s,B_{T-s}^x)\frac{B_{T-s}}{T-s}ds\right\}\right]^{1/4}\\
&=: I\cdot II\cdot III.
\end{aligned}$$

where $W_t := \int_0^t \frac{B_{T-s}}{T-s}ds + B_{T-t} - B_T$ is a Brownian motion with respect to the filtration generated by $\hat{B}_\cdot$. For factor I, Hölder's inequality gives

$$\begin{aligned}
E\bigg[\exp\bigg\{-2\lambda\int_0^t f(s,B_s^x)dB_s\bigg\}\bigg] &\le\\
&\le E\left[\mathcal{E}\left(\int_0^t(-4\lambda f(s,B_s^x))dB_s\right)\right]^{1/2} E\left[\exp\left\{\int_0^t(8\lambda^2 f^2(s,B_s^x))ds\right\}\right]^{1/2}\\
&= E\left[\exp\left\{\int_0^t(8\lambda^2 f^2(s,B_s^x))ds\right\}\right]^{1/2}\\
&\le C,
\end{aligned}$$

where $C>0$ is independent of x since f is bounded. Analogously, we obtain a bound for II. Finally, III follows from

$$E\left[\exp\left\{k\int_0^T\frac{|B_s|}{s}ds\right\}\right] < \infty \tag{6.71}$$

for any $k\in\mathbb{R}$, see Lemma 6.32 below. □

Lemma 6.32. *Let B be a one-dimensional Brownian motion on $[0,T]$. Then for any integer $p\ge 1$ and $0\le\varepsilon<1/(4p)$*

$$E\left[\left|\int_0^T\frac{|B_u|^{1+\varepsilon}}{u^{1+\varepsilon}}du\right|^p\right] < \infty. \tag{6.72}$$

Proof. Indeed,

$$
\begin{aligned}
E\left[\left|\int_0^T \frac{|B_u|^{1+\varepsilon}}{u^{1+\varepsilon}}du\right|^p\right] &\leq E\left[\left(\sup_{u\in[0,T]}|B_u|^{\varepsilon}\right)^p\left|\int_0^T \frac{|B_u|}{u^{1+\varepsilon}}du\right|^p\right]\\
&\leq E\left[\sup_{u\in[0,T]}|B_u|^{2p\varepsilon}\right]^{1/2} E\left[\left|\int_0^T \frac{|B_u|}{u^{1+\varepsilon}}du\right|^{2p}\right]^{1/2}\\
&\leq CE\left[\left|\int_0^T \frac{|B_u|}{u^{1+\varepsilon}}du\right|^{2p}\right]^{1/2}
\end{aligned}
$$

for a positive constant $C > 0$. Now, set $d := 2p$ then we may write

$$
\begin{aligned}
E\left[\left|\int_0^T \frac{|B_u|}{u^{1+\varepsilon}}du\right|^{2p}\right] &= \int_0^T \overset{d)}{\cdots} \int_0^T \frac{E\left[|B_{u_1}|\cdots|B_{u_d}|\right]}{u_1^{1+\varepsilon}\cdots u_d^{1+\varepsilon}}du_1\cdots du_d\\
&= d!\int_{0<u_1<\cdots<u_d<T}\frac{E\left[|B_{u_1}|\cdots|B_{u_d}|\right]}{u_1^{1+\varepsilon}\cdots u_d^{1+\varepsilon}}du_1\cdots du_d
\end{aligned}
\tag{6.73}
$$

where the last equality follows from the fact that the integrand is a symmetric function.

Then for a centered random Gaussian vector $(Z_1, \ldots, Z_d)$ with covariances $Cov(Z_i, Z_j) = \sigma_{i,j}$, $i, j = 1, \ldots, d$ we have the following estimate that can be found in [76, Theorem 1]

$$
E[|Z_1\cdots Z_d|] \leq \left(\sum_{\pi\in S_d}\prod_{j=1}^d \sigma_{j,\pi(j)}\right)^{1/2}
$$

where S_d denotes the set of permutations of $(1, \ldots, d)$. Applying the above inequality to the integral in (6.73)

$$
\begin{aligned}
&\int_{0<u_1<\cdots<u_d<T}\frac{E\left[|B_{u_1}|\cdots|B_{u_d}|\right]}{u_1^{1+\varepsilon}\cdots u_d^{1+\varepsilon}}du_1\cdots du_d \leq\\
&\qquad\leq \sum_{\pi\in S_d}\int_{0<u_1<\cdots<u_d<T}\prod_{j=1}^d\left(\frac{u_j\wedge u_{\pi(j)}}{u_j^{1+\varepsilon}u_{\pi(j)}^{1+\varepsilon}}\right)^{1/2}du_1\cdots du_d.
\end{aligned}
$$

Given a permutation $\pi \in S_d$ we have that, if $0 < u_1 < u_2 < \cdots u_d < T$ then

$$
\prod_{j=1}^d\left(\frac{u_j\wedge u_{\pi(j)}}{u_j^{1+\varepsilon}u_{\pi(j)}^{1+\varepsilon}}\right)^{1/2} = \frac{u_1^{\alpha_1/2}\cdots u_d^{\alpha_d/2}}{u_1^{1+\varepsilon}\cdots u_d^{1+\varepsilon}}
$$

where the α_i's, depend on π and have the property that $\sum_{i=1}^d \alpha_i = d$ and $\alpha_i \in \{0, 1, 2\}$ for all $i = 1, \ldots, d$. Moreover, observe that $\alpha_1 \geq 1$ independently of π since $u_1 \wedge u_{\pi(1)} = u_1$ for all $\pi \in S_d$. So, if we now integrate iteratively we obtain

$$
\int_{0<u_1<\cdots<u_d<T}\frac{E\left[|B_{u_1}|\cdots|B_{u_d}|\right]}{u_1^{1+\varepsilon}\cdots u_d^{1+\varepsilon}}du_1\cdots du_d \leq \sum_{\pi\in S_d}\frac{1}{\prod_{j=1}^d\left(\frac{1}{2}\sum_{i=1}^j\alpha_i - j\varepsilon\right)}T^{d(\frac{1}{2}-\varepsilon)}
$$

if, and only if $\frac{1}{2}\sum_{i=1}^{j}\alpha_i - j\varepsilon > 0$ for all $j = 1, \dots, d$ which holds by just observing that

$$\frac{1}{2}\sum_{i=1}^{j}\alpha_i > \frac{\alpha_1}{2} \geq d\frac{1}{2d} \geq j\frac{1}{2d}$$

for every $j = 1, \dots, d$ where we used $\alpha_1 \geq 1$. So it suffices to take $\varepsilon \geq 0$ such that $\varepsilon < \frac{1}{2d}$. □

6.A.2 Proof of Theorem 6.14

We now develop the proof of Theorem 6.14 according to the four-step scheme outlined in Section 6.3. In order to construct a weak solution of (6.5) in the first step, let $(\Omega, \mathcal{F}, \widetilde{P})$ be some given probability space which carries a Brownian motion $\widetilde{B}$, and put $X_t^x := \widetilde{B}_t + x$, $t \in [0,T]$. As we already noted in Remark 6.8, it is well-known, see e.g. [62, Corollary 5.16], that for sublinear coefficients b the Radon-Nikodym derivative $\frac{dP}{d\widetilde{P}} := \mathcal{E}\left(\int_0^T b(u, X_u^x)d\widetilde{B}_u\right)$ defines an equivalent probability measure P under which the process

$$B_t := X_t^x - x - \int_0^t b(s, X_s^x)ds,\ t \in [0,T], \tag{6.74}$$

is a Brownian motion on $(\Omega, \mathcal{F}, P)$. Hence, because of (6.74), the pair (X^x, B) is a weak solution of (6.5) on $(\Omega, \mathcal{F}, P)$. The stochastic basis that we operate on in the following is now given by the filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\in[0,T]})$, which carries the weak solution (X^x, B) of (6.5), where $\{\mathcal{F}_t\}_{t\in[0,T]}$ denotes the filtration generated by B_t, $t \in [0,T]$, augmented by the P-null sets.

Next, we prove that for given $t \in [0,T]$ the sequence of strong solutions $\{X_t^{n,x}\}_{n\geq 1}$ of the SDE's (6.24) with regular coefficients b_n from (6.23) converges weakly in $L^2(\Omega; \mathcal{F}_t)$ to $E[X_t^x|\mathcal{F}_t]$.

Lemma 6.33. *Let $b_n : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a sequence of functions approximating b a.e. as in (6.23) and $X_t^{n,x}$ the corresponding strong solutions to* (6.24)*, $n \geq 1$. Then for every $t \in [0,T]$ and function $\varphi \in L_w^{2p}(\mathbb{R})$ where the space $L_w^{2p}(\mathbb{R})$ is defined as in* (6.37) *with p being the conjugate exponent of $1+\varepsilon$, $\varepsilon > 0$ from Lemma 6.29, we have*

$$\varphi(X_t^{n,x}) \xrightarrow{n\to\infty} E[\varphi(X_t^x)|\mathcal{F}_t]$$

weakly in $L^2(\Omega; \mathcal{F}_t)$.

Proof. First of all, we shall see that $\varphi(X_t^{n,x}), E[\varphi(X_t^x)|\mathcal{F}_t] \in L^2(\Omega; \mathcal{F}_t)$, $n \geq 0$. Indeed, Girsanov's theorem, Remark 6.30 and the fact that $\varphi \in L_w^{2p}(\mathbb{R})$ imply that for some constant $C_\varepsilon > 0$ with $\varepsilon > 0$ small enough we have

$$\sup_{n\geq 0} E[|\varphi(X_t^{n,x})|^2] \leq C_\varepsilon E[|\varphi(x+B_t)|^{2\frac{1+\varepsilon}{\varepsilon}}]^{\frac{\varepsilon}{1+\varepsilon}} = C_\varepsilon \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} |\varphi(x+z)|^{2\frac{1+\varepsilon}{\varepsilon}} e^{-\frac{|z|^2}{2T}}dz < \infty. \tag{6.75}$$

To show that

$$E\left[\varphi(X_t^{n,x})Z\right] \xrightarrow{n\to\infty} E[E[\varphi(X_t^x)|\mathcal{F}_t]Z]$$

for any $Z \in L^2(\Omega; \mathcal{F}_t)$ it suffices to show

$$\mathcal{W}(X_t^{n,x})(f) \xrightarrow{n\to\infty} \mathcal{W}(E[X_t^x|\mathcal{F}_t])(f)$$

for every $f \in L^2([0,T])$

Indeed, by Girsanov's theorem we can write

$$\begin{aligned}
E\Bigg[\Big(\varphi(X_t^{n,x}) - E[\varphi(X_t^x)|\mathcal{F}_t]\Big)\mathcal{E}\left(\int_0^T f(u)dB_u\right)\Bigg] = \\
&= E\left[\varphi(B_t^x)\left(\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right) - \mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\right)\right] \\
&= E\Bigg[\varphi(B_t^x)\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right) \\
&\quad \times \left(\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right)\Big/\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right) - 1\right)\Bigg]
\end{aligned}$$

Then, using inequality $|e^x - 1| \le |x|(e^x+1)$ we have

$$\begin{aligned}
E\Bigg[\Big(\varphi(X_t^{n,x}) - E[\varphi(X_t^x)|\mathcal{F}_t]\Big)\mathcal{E}\left(\int_0^T f(u)dB_u\right)\Bigg] \\
&\le E\left[|\varphi(B_t^x)|\,|U_n|\,\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right)\right] \\
&\quad + E\left[|\varphi(B_t^x)|\,|U_n|\,\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\right] \\
&:= I_n + II_n
\end{aligned}$$

where

$$U_n := \int_0^T (b_n(u,B_u^x) - b(u,B_u^x))dB_u - \frac{1}{2}\int_0^T [(b_n(u,B_u^x)+f(u))^2 - (b(u,B_u^x)+f(u))^2]du.$$

For the term I_n, Hölder's inequality with exponents $p = \frac{1+\varepsilon}{\varepsilon}$ and $q = 1+\varepsilon$ and then again for $p = q = 2$ yields

$$\begin{aligned}
I_n &\le E\left[|\varphi(B_t^x)U_n|^{\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{1+\varepsilon}} E\left[\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right)^{1+\varepsilon}\right]^{\frac{1}{1+\varepsilon}} \\
&\le E\left[|\varphi(B_t^x)|^{2\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} E\left[|U_n|^{2\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} E\left[\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))\,dB_u\right)^{1+\varepsilon}\right]^{\frac{1}{1+\varepsilon}} \\
&=: I^1 \cdot I_n^2 \cdot I_n^3,
\end{aligned}$$

where I^1, I_n^2 and I_n^3 are the respective factors above and $\varepsilon > 0$ is such that I_n^3 is bounded uniformly in $n \geq 0$ (see Remark 6.30). We can then control the first factor I^1 due to the fact that $\varphi \in L_w^{2p}(\mathbb{R})$ as it is shown in (6.75).

Finally, for the second factor I_n^2 define $p := 2\frac{1+\varepsilon}{\varepsilon}$. Then using Minkowski's inequality, Burkholder-Davis-Gundy's inequality and Hölder's inequality we can write

$$\begin{aligned}
(I_n^2)^p &= E\Bigg[\Bigg|\int_0^T (b_n(u,B_u^x) - b(u,B_u^x))dB_u \\
&\quad - \frac{1}{2}\int_0^T [(b_n(u,B_u^x) + f(u))^2 - (b(u,B_u^x) + f(u))^2]du\Bigg|^p\Bigg] \\
&\leq 2^{p-1}E\left[\left|\int_0^T (b_n(u,B_u^x) - b(u,B_u^x))dB_u\right|^p\right] \\
&\quad + 2^{p-2}E\left[\left|\int_0^T [(b_n(u,B_u^x) + f(u))^2 - (b(u,B_u^x) + f(u))^2]du\right|^p\right] \\
&\lesssim 2^{p-1}E\left[\left(\int_0^T |b_n(u,B_u^x) - b(u,B_u^x)|^2\,du\right)^{p/2}\right] \\
&\quad + 2^{p-2}T^{p-1}\int_0^T E\left[\left|(b_n(u,B_u^x) + f(u))^2 - (b(u,B_u^x) + f(u))^2\right|^{2p}\right] du \\
&\lesssim 2^{p-1}T^{p/2-1}\int_0^T E\left[|b_n(u,B_u^x) - b(u,B_u^x)|^p\right] du \\
&\quad + 2^{p-2}T^{p-1}\int_0^T E\left[\left|(b_n(u,B_u^x) + f(u))^2 - (b(u,B_u^x) + f(u))^2\right|^{2p}\right] du
\end{aligned}$$

and by dominated convergence we obtain $I_n^2 \to 0$ as $n \to \infty$. Similarly, we obtain the result for II_n. □

We now turn to the third step of our scheme to prove Theorem 6.14. The next theorem gives the $L^2(\Omega;\mathcal{F}_t)$-convergence of the sequence of strong solutions $X_t^{n,x}$ to the limit $E[X_t^x|\mathcal{F}_t]$ which, in addition, is Malliavin differentiable. The technique used in this result is the compactness criterion given in Proposition 6.6 due to [30].

Theorem 6.34. *Let $b_n : [0,T]\times\mathbb{R} \to \mathbb{R}$, $n \geq 1$, be as in (6.23) and $X_\cdot^{n,x}$ the corresponding strong solutions to (6.24). Then, for each $t \in [0,T]$*

$$X_t^{n,x} \xrightarrow{L^2(\Omega;\mathcal{F}_t)} E[X_t^x|\mathcal{F}_t] \tag{6.76}$$

as $n \to \infty$. Moreover, the right-hand side of (6.76) is Malliavin differentiable.

Proof. The main step is to show relative compactness of $\{X_t^{n,x}\}_{n\geq 1}$ by applying Proposition 6.6. Let $t \in [0,T]$, $0 \leq s \leq s' \leq t$ and a compact set $K \subset \mathbb{R}$ be given. Using the explicit representation introduced in (6.25), Girsanov's theorem, the mean-value theorem, Hölder's inequality with exponent $1+\varepsilon$ for a sufficiently small $\varepsilon > 0$ and Cauchy-Schwarz inequality successively we obtain

$$
\begin{aligned}
&E\big[(D_s X_t^{n,x} - D_{s'} X_t^{n,x})^2\big] = \\
&= E\left[\exp\left\{2\int_{s'}^t b_n'(u,B_u^x)du\right\}\left(\exp\left\{\int_s^{s'} b_n'(u,B_u^x)du\right\}-1\right)^2 \mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)\right] \\
&\le E\Bigg[\exp\left\{2\int_{s'}^t b_n'(u,B_u^x)du\right\}\left(\sup_{0\le\alpha\le1}\exp\left\{\alpha\int_s^{s'} b_n'(u,B_u^x)du\right\}\right)^2 \\
&\quad\times\left(\int_s^{s'} b_n'(u,B_u^x)du\right)^2 \mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)\Bigg] \\
&\le E\Bigg[\exp\left\{2\frac{1+\varepsilon}{\varepsilon}\int_{s'}^t b_n'(u,B_u^x)du\right\}\sup_{0\le\alpha\le1}\exp\left\{2\frac{1+\varepsilon}{\varepsilon}\alpha\int_s^{s'} b_n'(u,B_u^x)du\right\} \\
&\quad\times\left|\int_s^{s'} b_n'(u,B_u^x)du\right|^{2\frac{1+\varepsilon}{\varepsilon}}\Bigg]^{\frac{\varepsilon}{1+\varepsilon}} E\left[\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)^{1+\varepsilon}\right]^{\frac{1}{1+\varepsilon}} \\
&\le E\left[\exp\left\{4\frac{1+\varepsilon}{\varepsilon}\int_{s'}^t(\tilde b_n'(u,B_u^x)+\hat b'(u,B_u^x))du\right\}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} \\
&\quad\times E\left[\sup_{0\le\alpha\le1}\exp\left\{8\frac{1+\varepsilon}{\varepsilon}\alpha\int_s^{s'}(\tilde b_n'(u,B_u^x)+\hat b'(u,B_u^x))du\right\}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} \\
&\quad\times E\left[\left|\int_s^{s'}(\tilde b_n'(u,B_u^x)+\hat b'(u,B_u^x))du\right|^{8\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} E\left[\mathcal{E}\left(\int_0^T b_n(u,B_u^x)dB_u\right)^{1+\varepsilon}\right]^{\frac{1}{1+\varepsilon}} \\
&=: I_n^1\cdot I_n^2\cdot I_n^3\cdot I_n^4,
\end{aligned}
$$

where I_n^1, I_n^2, I_n^3 and I_n^4 denote the respective factors shown above.

Here, by Remark 6.30, $\varepsilon > 0$ is chosen such that

$$\sup_{x\in K}\sup_{n\ge0} I_n^4 < \infty.$$

For I_n^1 and I_n^2 we use Cauchy-Schwarz inequality and the fact that $\hat b'$ is bounded and get

$$I_n^1 \lesssim E\left[\exp\left\{4\frac{1+\varepsilon}{\varepsilon}\int_{s'}^t \tilde b_n'(u,B_u^x)du\right\}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} =: II_n^1$$

and

$$I_n^2 \lesssim E\left[\sup_{0\le\alpha\le1}\exp\left\{8\frac{1+\varepsilon}{\varepsilon}\alpha\int_s^{s'}\tilde b_n'(u,B_u^x)du\right\}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} =: II_n^2.$$

For I_n^3, Minkowski's inequality and the boundedness of $\hat{b}'$ give

$$\begin{aligned}
I_n^3 &\leq E\left[\left|\int_s^{s'} \tilde{b}_n'(u,B_u^x)du\right|^{8\frac{1+\varepsilon}{\varepsilon}} + \left|\int_s^{s'} \hat{b}'(u,B_u^x)du\right|^{8\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} \\
&\lesssim E\left[\left|\int_s^{s'} \tilde{b}_n'(u,B_u^x)du\right|^{8\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} + \|\hat{b}'\|_\infty^2 T|s'-s| \\
&\leq II_n^3 + \|\hat{b}'\|_\infty^2 T|s'-s|.
\end{aligned}$$

Now we want to get rid of the derivatives $\tilde{b}_n'$ in II_n^1, II_n^2 and II_n^3. In order to do so, we use integration with respect to the local time of the Brownian motion, see Theorem 6.12 in the Section 6.2 or e.g. [39] for more information about local-time integration. We obtain

$$\begin{aligned}
E\big[(D_sX_t^{n,x} - D_{s'}X_t^{n,x})^2\big] \lesssim\ & E\left[\exp\left\{-4\frac{1+\varepsilon}{\varepsilon}\int_{s'}^t\int_{\mathbb{R}} \tilde{b}_n(u,y)L^{B^x}(du,dy)\right\}\right]^{\frac{\varepsilon}{2(1+\varepsilon)}} \\
&\times E\left[\sup_{0\leq\alpha\leq 1}\exp\left\{-8\frac{1+\varepsilon}{\varepsilon}\alpha\int_s^{s'}\int_{\mathbb{R}} \tilde{b}_n(u,x)L^{B^x}(du,dy)\right\}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} \\
&\times\left(E\left[\left|\int_s^{s'}\int_{\mathbb{R}} \tilde{b}_n(u,x)L^{B^x}(du,dy)\right|^{8\frac{1+\varepsilon}{\varepsilon}}\right]^{\frac{\varepsilon}{4(1+\varepsilon)}} + \|\hat{b}'\||s'-s|\right).
\end{aligned}$$

Observe that factors II_n^1 and II_n^2 can be controlled uniformly in $n \geq 1$ and $x \in K$ by virtue of Lemma 6.31. Now, denote $p_\varepsilon := 4\frac{1+\varepsilon}{\varepsilon}$. Then for factor II_n^3 we use representation (6.20) from Theorem 6.12 in connection with (6.19) in Section 6.2 and apply Minkowski's inequality, Burkholder-Davis-Gundy's inequality and Hölder's inequality with exponent $(\varepsilon'+2)/\varepsilon'$ for a suitable $\varepsilon' > 0$ in order to obtain

$$\begin{aligned}
II_n^3 \leq\ & E\Bigg[\Bigg|\int_s^{s'} \tilde{b}_n(u,B_u^x)dB_u - \int_{T-s'}^{T-s} \tilde{b}_n(T-u,\hat{B}_u^x)dW_u \\
&\qquad + \int_{T-s'}^{T-s} \tilde{b}_n(T-u,\hat{B}_u^x)\frac{\hat{B}_u}{T-u}du\Bigg|^{2p_\varepsilon}\Bigg]^{1/p_\varepsilon} \\
\lesssim\ & E\left[\left(\int_s^{s'} |\tilde{b}_n(u,B_u^x)|^2 du\right)^{p_\varepsilon}\right]^{1/p_\varepsilon} + E\left[\left(\int_{T-s'}^{T-s} |\tilde{b}_n(T-u,\hat{B}_u^x)|^2 du\right)^{p_\varepsilon}\right]^{1/p_\varepsilon} \\
&+ E\left[\left|\int_{T-s'}^{T-s} \tilde{b}_n(T-u,\hat{B}_u^x)\frac{\hat{B}_u}{T-u}du\right|^{2p_\varepsilon}\right]^{1/p_\varepsilon}
\end{aligned}$$

$$
\begin{aligned}
&\lesssim |s'-s|^{\varepsilon'/(\varepsilon'+2)} E\left[\left(\int_s^{s'} |\tilde{b}_n(u, B_u^x)|^{\varepsilon'+2} du\right)^{\frac{2p_\varepsilon}{\varepsilon'+2}}\right]^{1/p_\varepsilon} \\
&+ |s'-s|^{\varepsilon'/(\varepsilon'+2)} E\left[\left(\int_{T-s'}^{T-s} |\tilde{b}_n(T-u, \hat{B}_u^x)|^{\varepsilon'+2} du\right)^{\frac{2p_\varepsilon}{\varepsilon'+2}}\right]^{1/p_\varepsilon} \\
&+ |s'-s|^{2\varepsilon'/(\varepsilon'+2)} E\left[\left|\int_{T-s'}^{T-s} \left|\tilde{b}_n(T-u, \hat{B}_u^x)\frac{\hat{B}_u}{T-u}\right|^{(\varepsilon'+2)/2} du\right|^{\frac{4p_\varepsilon}{\varepsilon'+2}}\right]^{1/p_\varepsilon}.
\end{aligned}
$$

The last expectation is bounded by taking $\varepsilon' < \frac{2}{8p_\varepsilon - 1}$ and applying Lemma 6.32.

Altogether, we can find a constant $C > 0$ such that

$$
\sup_{x\in K} \sup_{n\geq 1} E\left[(D_{s'}X_t^{n,x} - D_sX_t^{n,x})^2\right] \leq C|s'-s|^{\varepsilon'/(\varepsilon'+2)} \tag{6.77}
$$

for $0 \leq s' \leq s \leq t$ where $0 < \varepsilon'/(\varepsilon'+2) < 1$.

Similarly, one also obtains

$$
\sup_{x\in K} \sup_{0\leq s\leq t} \sup_{n\geq 1} E\left[(D_sX_t^{n,x})^2\right] \leq C \tag{6.78}
$$

for a constant $C > 0$.

Then (6.75) with $\varphi = id$, (6.77), (6.78) together with Proposition 6.6 imply that the set $\{X_t^{n,x}\}_{n\geq 1}$ is relatively compact in $L^2(\Omega; \mathcal{F}_t)$. Since the sequence of solutions $X_t^{n,x}$ also converges weakly to $E[X_t^x|\mathcal{F}_t]$ due to Lemma 6.33 with $\varphi = id$, by uniqueness of the limit we have that

$$
X_t^{n_k,x} \xrightarrow{L^2(\Omega;\mathcal{F}_t)} E[X_t^x|\mathcal{F}_t]
$$

for a subsequence n_k, $k \geq 0$.

In fact, one observes that the $L^2(\Omega; \mathcal{F}_t)$-convergence holds for the whole sequence. Indeed, assume by contradiction, that there exists a subsequence n_j, $j \geq 0$, such that there is an $\varepsilon > 0$ with $E[|X_t^{n_j,x} - X_t^x|^2] > \varepsilon$ for all $j \geq 0$. Then $\{b_{n_j}\}_{j\geq 0}$ is a sequence of approximating coefficients as required in (6.23). Thus, by the previous results there exists a subsequence n_{j_m}, $m \geq 0$, such that $X^{n_{j_m},x} \to X^x$ in $L^2(\Omega; \mathcal{F}_t)$, which gives rise to a contradiction.

Moreover, since the sequence of Malliavin derivatives $\{D_sX_t^{n,x}\}_{n\geq 1}$ is bounded uniformly in n in the $L^2([0,T]\times\Omega)$-norm because of (6.78), we also have that the limit $E[X_t^x|\mathcal{F}_t]$ is Malliavin differentiable, see for instance [90, Lemma 1.2.3]. □

Remark 6.35. *Note that we have proved the estimates (6.77) and (6.78) uniformly in $x \in K$ for a compact set K even though this is not needed to apply Proposition 6.6. We will, however, use this uniform bounds later on in the proofs of Lemma 6.36 and Theorem 6.17.*

We are now ready to complete the proof of Theorem 6.14 by use of the previous steps.

Proof of Theorem 6.14. It remains to prove that X_t^x is $\mathcal{F}_t$-measurable for every $t \in [0,T]$ and by Remark 6.3 it then follows that there exists a strong solution in the usual sense that is Malliavin differentiable. Indeed, let φ be a continuous bounded function, then by Theorem 6.34 we have, for a subsequence n_k, $k \geq 0$, that

$$\varphi(X_t^{n_k,x}) \to \varphi(E[X_t^x|\mathcal{F}_t]), \;\; P-a.s.$$

as $k \to \infty$.

On the other side, by Lemma 6.33 we also have

$$\varphi(X_t^{n,x}) \to E\left[\varphi(X_t^x)|\mathcal{F}_t\right]$$

weakly in $L^2(\Omega;\mathcal{F}_t)$. By the uniqueness of the limit we immediately have

$$\varphi\left(E[X_t^x|\mathcal{F}_t]\right) = E\left[\varphi(X_t^x)|\mathcal{F}_t\right], \;\; P-a.s.$$

for all continuous, bounded functions φ, which implies that X_t^x is $\mathcal{F}_t$-measurable for every $t \in [0,T]$.

To show uniqueness, assume that we have two strong solutions X^x and Y^x to the SDE (6.5). Then using the Cameron-Martin formula shows that

$$\mathcal{W}(X_t^x)(h) = E[X_t^x(h)],$$

for $h \in L^2([0,T])$ where we recall that $\mathcal{W}(X_t^x)(h)$ denotes the Wiener transform, and the process $X_t^x(h), 0 \leq t \leq T$ satisfies the SDE

$$dX_t^x(h) = (b(t, X_t^x(h)) + h(t))dt + d\widehat{B}_t, \;\; X_0^x(h) = x \tag{6.79}$$

for a Brownian motion $\widehat{B}_t, 0 \leq t \leq T$. In the same way, the process $Y_t^x(h), 0 \leq t \leq T$ solves (6.79). On the other hand, it follows from the linear growth of the drift coefficient b that $X_t^x(h)$ and $Y_t^x(h), 0 \leq t \leq T$, are unique in law (see e.g. Proposition 3.10 in [62]). Hence

$$\mathcal{W}(X_t^x)(h) = \mathcal{W}(Y_t^x)(h)$$

for all t, h. Thus X^x and Y^x are indistinguishable. □

6.A.3 Proof of Proposition 6.15:

By equation (6.25) and formula (6.20), we can write for regular coefficients b_n

$$D_s X_t^{n,x} = \exp\left\{-\int_s^t \int_{\mathbb{R}} b_n(u,y) L^{X^{n,x}}(du,dy)\right\}.$$

Then, since $X_t^{n,x}, n \geq 0$ is relatively compact in $L^2(\Omega;\mathcal{F}_t)$ and $\|D_s X_t^{n,x}\|_{L^2([0,T]\times\Omega)}$ is bounded uniformly in $n \geq 0$ due to the proof of Theorem 6.34 we know that the sequence $D_s X_t^{n,x}$, $n \geq 0$ converges weakly to $D_s X_t^x$ in $L^2([0,T]\times\Omega)$, see [90, Lemma 1.2.3]. Therefore, it is enough to check that our candidate is the weak limit. So we must prove that

$$\left\langle \mathcal{W}\Big(\exp\Big\{-\int_\cdot^t\int_{\mathbb{R}} b_n(u,y)L^{X^{n,x}}(du,dy)\Big\} - \exp\Big\{-\int_\cdot^t\int_{\mathbb{R}} b(u,y)L^{X^x}(du,dy)\Big\}\Big)(f), g\right\rangle_{L^2([0,T])}$$

converges to 0 as $n\to\infty$ for every $f\in L^2([0,T])$ and $g\in C_0^\infty([0,T])$. It suffices to show that the Wiener transform goes to zero.

Then, as we did for Lemma 6.33, using Girsanov's theorem we have

$$\begin{aligned}
&\Bigg|E\Bigg[\mathcal{E}\left(\int_0^T f(u)dB_u\right)\Bigg(\exp\Big\{-\int_s^t\int_{\mathbb{R}} b_n(u,y)L^{X^{n,x}}(du,dy)\Big\}\\
&\qquad\qquad -\exp\Big\{-\int_s^t\int_{\mathbb{R}} b(u,y)L^{X^x}(du,dy)\Big\}\Bigg)\Bigg]\Bigg|\\
&=\Bigg|E\Bigg[\exp\Big\{-\int_s^t\int_{\mathbb{R}} b_n(u,y)L^{B^x}(du,dy)\Big\}\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right)\\
&\quad -\exp\Big\{-\int_s^t\int_{\mathbb{R}} b(u,y)L^{B^x}(du,dy)\Big\}\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\Bigg]\Bigg|\\
&\le\Bigg|E\Bigg[\Bigg(\exp\Big\{-\int_s^t\int_{\mathbb{R}} \tilde b_n(u,y)L^{B^x}(du,dy)\Big\}-\exp\Big\{-\int_s^t\int_{\mathbb{R}} \tilde b(u,y)L^{B^x}(du,dy)\Big\}\Bigg)\\
&\quad\times\exp\Big\{\int_s^t \hat b'(u,B_u^x)du\Big\}\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\Bigg]\Bigg|\\
&\quad+\Bigg|E\Bigg[\Bigg(\mathcal{E}\left(\int_0^T (b_n(u,B_u^x)+f(u))dB_u\right)-\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\Bigg)\\
&\quad\times\exp\Big\{-\int_s^t\int_{\mathbb{R}} \tilde b_n(u,y)L^{B^x}(du,dy)\Big\}\exp\Big\{\int_s^t \hat b'(u,B_u^x)du\Big\}\Bigg]\Bigg|\\
&=: I_n + II_n.
\end{aligned}$$

For term I_n we define $p := \frac{1+\varepsilon}{\varepsilon}$ for a suitable $\varepsilon>0$ and then apply Hölder's inequality with exponent $1+\varepsilon$ on the stochastic exponential. Then we apply Cauchy-Schwarz inequality and bound the factor with $\|\hat b'\|_\infty$, and finally we use inequality $|e^x-1|\le|x|(e^x+1)$. As a result we obtain

$$\begin{aligned}
I_n =&\Bigg|E\Bigg[\exp\Big\{-\int_s^t\int_{\mathbb{R}} \tilde b(u,y)L^{B^x}(du,dy)\Big\}\\
&\times\Bigg(\exp\Big\{-\int_s^t\int_{\mathbb{R}} (\tilde b_n(u,y)-\tilde b(u,y))L^{B^x}(du,dy)\Big\}-1\Bigg)\\
&\times\exp\Big\{\int_s^t \hat b'(u,B_u^x)du\Big\}\mathcal{E}\left(\int_0^T (b(u,B_u^x)+f(u))dB_u\right)\Bigg]\Bigg|
\end{aligned}$$

$$\lesssim E\Bigg[\exp\left\{-2p\int_s^t\int_{\mathbb{R}}\tilde{b}(u,y)L^{B^x}(du,dy)\right\}$$

$$\times\left|\left(\exp\left\{-\int_s^t\int_{\mathbb{R}}(\tilde{b}_n(u,y)-\tilde{b}(u,y))L^{B^x}(du,dy)\right\}-1\right)^{2p}\right|\Bigg]^{1/(2p)}$$

$$\times E\left[\mathcal{E}\left(\int_0^T(b(u,B_u^x)+f(u))dB_u\right)^{1+\varepsilon}\right]^{1/(1+\varepsilon)}$$

$$\lesssim E\Bigg[\left|\int_s^t\int_{\mathbb{R}}(\tilde{b}_n(u,y)-\tilde{b}(u,y))L^{B^x}(du,dy)\right|^{2p}\Bigg(\exp\left\{-\int_s^t\int_{\mathbb{R}}\tilde{b}_n(u,y)L^{B^x}(du,dy)\right\}$$

$$+\exp\left\{-\int_s^t\int_{\mathbb{R}}\tilde{b}(u,y)L^{B^x}(du,dy)\right\}\Bigg)^{2p}\Bigg]^{1/(2p)}$$

where in the last inequality we choose $\varepsilon > 0$ small enough so that the stochastic exponential is bounded due to Lemma 6.29. Then Minkowski's inequality gives

$$\begin{aligned}(I_n)^{2p}\lesssim\ & E\left[|V_n|^{2p}\exp\left\{-2p\int_s^t\int_{\mathbb{R}}\tilde{b}_n(u,y)L^{B^x}(du,dy)\right\}\right]\\ &+E\left[|V_n|^{2p}\exp\left\{-2p\int_s^t\int_{\mathbb{R}}\tilde{b}(u,y)L^{B^x}(du,dy)\right\}\right]\end{aligned}\tag{6.80}$$

where

$$V_n:=\int_s^t\int_{\mathbb{R}}(\tilde{b}_n(u,y)-\tilde{b}(u,y))L^{B^x}(du,dy).$$

Then Cauchy-Schwarz inequality and Lemma 6.31 give

$$\begin{aligned}&E\left[|V_n|^{2p}\exp\left\{-2p\int_s^t\int_{\mathbb{R}}\tilde{b}_n(u,y)L^{B^x}(du,dy)\right\}\right]\le\\ &\qquad\le E\left[|V_n|^{4p}\right]^{1/2}E\left[\exp\left\{-4p\int_s^t\int_{\mathbb{R}}\tilde{b}_n(u,y)L^{B^x}(du,dy)\right\}\right]^{1/2}\\ &\qquad\lesssim E\left[|V_n|^{4p}\right]^{1/2}.\end{aligned}\tag{6.81}$$

Finally, using representation (6.19) in the Section 6.2, Minkowski's inequality, Burkholder-Davis-Gundy's inequality in the first two terms and Hölder's inequality in the last term we obtain

$$E\big[|V_n|^p\big]=E\Bigg[\Bigg|\int_s^t(\tilde{b}_n(u,B_u^x)-\tilde{b}(u,B_u^x))dB_u+\int_{T-t}^{T-s}(\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x))dW_u$$

$$-\int_{T-t}^{T-s}(\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x))\frac{\hat{B}_u}{T-u}du\Bigg|^p\Bigg]$$

$$
\begin{aligned}
&\leq E\left[\left|\int_s^t (\tilde{b}_n(u,B_u^x)-\tilde{b}(u,B_u^x))dB_u\right|^p + E\left[\left|\int_{T-t}^{T-s} (\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x))dW_u\right|^p\right]\right. \\
&+ E\left[\left|\int_{T-t}^{T-s} (\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x))\frac{\hat{B}_u}{T-u}du\right|^p\right] \\
&\leq E\left[\left[\int_s^t |\tilde{b}_n(u,B_u^x)-\tilde{b}(u,B_u^x)|^2 du\right]^{p/2} + E\left[\int_{T-t}^{T-s} |\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x)|^2 du\right]^{p/2}\right] \\
&+ E\left[\left|\int_{T-t}^{T-s} (\tilde{b}_n(T-u,\hat{B}_u^x)-\tilde{b}(T-u,\hat{B}_u^x))\frac{\hat{B}_u}{T-u}du\right|^p\right].
\end{aligned}
$$

By dominated convergence, all terms converge to zero as $n \to \infty$. In order to justify that the third term also converges to 0 one needs to use the estimate in Lemma 6.32. The second term in (6.80) is estimated in the same way. Similarly, one can also bound II_n.

Lemma 6.36. *Let $b_n : [0,T] \times \mathbb{R} \to \mathbb{R}$, $n \geq 0$ be as in (6.23) and $X_t^{n,x}$ the corresponding strong solutions with drift coefficients b_n. Then, for any compact subset $K \subset \mathbb{R}$ and $p \geq 1$*

$$
\sup_{n\geq 1}\sup_{x\in K}\sup_{t\in[0,T]} E\left[\left(\frac{\partial}{\partial x}X_t^{n,x}\right)^p\right] \leq C_{K,p}
$$

for a constant $C_{K,p} > 0$ depending on K and p. Here, $\frac{\partial}{\partial x}X_t^{n,x}$ is the first variation process of $X_t^{n,x}$, $n \geq 1$ (see Proposition 6.7).

Proof. The proof of this result relies on the proof of (6.78) in Theorem 6.34 by observing that $\frac{\partial}{\partial x}X_t^{n,x} = D_0 X_t^{n,x}$ by Proposition 6.7. Then following exactly the same steps as in Theorem 6.34 we see that all computations can be done for an arbitrary power $p \geq 1$. Finally, from the term II_n^1 in the proof of Theorem 6.34 one can see that $\sup_{n\geq 1}\sup_{x\in K}\sup_{t\in[0,T]} E\left[\left(\frac{\partial}{\partial x}X_t^{n,x}\right)^p\right] < \infty$. □

6.A.4 Proof of Proposition 6.16:

First, start observing that, for any given $p \geq 1$, we have

$$
\begin{aligned}
E\left[|X_t^{n,x}|^p\right] &\lesssim |x|^p + \int_0^t E\left[|\tilde{b}_n(u,X_u^{n,x})|^p\right]du + \int_0^t E\left[|\hat{b}(u,X_u^{n,x})|^p\right]du + E\left[|B_t|^p\right] \\
&\lesssim |x|^p + |t|^p + C\int_0^t E\left[|X_u^{n,x}|^p\right]du
\end{aligned}
$$

due to the uniform boundedness of $\tilde{b}_n$, the continuity of $\hat{b}$ and Hölder continuity of the Brownian motion. Then, Grönwall's inequality gives

$$
\sup_{n\geq 1} E\left[|X_t^{n,x}|^p\right] \leq C. \tag{6.82}
$$

Now, assume that $0 \leq s < t \leq T$. Then

$$\begin{aligned} X_t^{n,x} - X_s^{n,y} &= x - y + \int_0^t b_n(u, X_u^{n,x})du - \int_0^s b_n(u, X_u^{n,y})du + B_t - B_s \\ &= x - y + \int_s^t b_n(u, X_u^{n,x})du + \int_0^s (b_n(u, X_u^{n,x}) - b_n(u, X_u^{n,y}))du + B_t - B_s. \end{aligned}$$

Now since b_n has linear growth together with (6.82), the uniform boundedness of $\tilde{b}_n$ and Hölder continuity of the Brownian motion yield

$$E\left[|X_t^{n,x} - X_s^{n,y}|^2\right] \lesssim |x-y|^2 + |t-s| + E\left[\left|\int_0^s (b_n(u, X_u^{n,x}) - b_n(u, X_u^{n,y}))du\right|^2\right].$$

Then we use the fact that $X_t^{n,s,\cdot}$ is a stochastic flow of diffeomorphisms (see e.g. [70]), the mean value theorem and Lemma 6.36 in order to obtain

$$\begin{aligned} &E\left[\left|\int_0^s (b_n(u, X_u^{n,x}) - b_n(u, X_u^{n,y}))du\right|^2\right] \\ &= |x-y|^2 E\left[\left|\int_0^s \int_0^1 b_n'(u, X_u^{n,x+\tau(y-x)}) \frac{\partial}{\partial x} X_u^{n,x+\tau(y-x)}) d\tau du\right|^2\right] \\ &\leq C|x-y|^2 \int_0^1 E\left[\left|\int_0^s b_n'(u, X_u^{n,x+\tau(y-x)}) \frac{\partial}{\partial x} X_u^{n,x+\tau(y-x)}) du\right|^2\right] d\tau \\ &= C|x-y|^2 \int_0^1 E\left[\left|\frac{\partial}{\partial x} X_s^{n,x+\tau(y-x)} - (1-\tau)\right|^2\right] d\tau \\ &\leq C|x-y|^2 \sup_{\substack{s\in[0,T] \\ x\in K}} E\left[\left|\frac{\partial}{\partial x} X_s^{n,x}\right|^2\right] \\ &\leq C|x-y|^2. \end{aligned}$$

Altogether

$$E\left[|X_t^{n,x} - X_s^{n,y}|^2\right] \leq C\left(|t-s| + |x-y|^2\right)$$

for a finite constant $C > 0$ independent of n.

To conclude, we use Fatou's lemma applied to a subsequence and the fact that $X_t^{n,x} \to X_t^x$ in $L^2(\Omega)$ as $n \to \infty$ due to Theorem 6.34.

6.A.5 Proof of Theorem 6.17

First of all, observe that for any smooth function with compact support $\varphi \in C_0^\infty(U, \mathbb{R})$ and $t \in [0, T]$, the sequence of random variables

$$\langle X_t^n, \varphi \rangle := \int_U X_t^{n,x} \varphi(x) dx$$

converges weakly in $L^2(\Omega)$ to $\langle X_t, \varphi\rangle$ by using the Wiener transform following the same steps as in Lemma 6.33.

Then for all measurable sets $A \in \mathcal{F}$, $\varphi \in C_0^\infty(\mathbb{R})$ and using Cauchy-Schwarz inequality we have

$$E[\mathbf{1}_A\langle X_t^{n_k,x} - X_t^x, \varphi'\rangle] \leq \|\varphi'\|_{L^2(U)}|U|^{1/2}\left(\sup_{x\in\text{supp}(U)} E\left[\mathbf{1}_A(X_t^{n_k,x} - X_t^x)^2\right]\right)^{1/2} < \infty$$

where the last quantity is finite by Proposition 6.16. Then by Theorem 6.34 we see that

$$\lim_{k\to\infty} E[\mathbf{1}_A\langle X_t^{n_k,x} - X_t^x, \varphi'\rangle] = 0.$$

In addition, by virtue of Lemma 6.36 we have that

$$\sup_{n\geq 0} E\|X_t^{n,x}\|^2_{W^{1,2}(U)} < \infty,$$

that is $x \mapsto X_t^{n,x}$ is bounded in $L^2(\Omega, W^{1,2}(U))$. As a result, the sequence $X_t^{n,x}$ is weakly relatively compact in $L^2(\Omega, W^{1,2}(U))$, see e.g. [75, Theorem 10.44], and therefore there exists a subsequence n_k, $k \geq 0$ such that $X_t^{n_k,x}$ converges weakly to some element $Y_t \in L^2(\Omega, W^{1,2}(U))$ as $k \to \infty$. Let us denote by Y_t' the weak derivative of Y_t.

Then

$$E[\mathbf{1}_A\langle X_t^x, \varphi'\rangle] = \lim_{k\to\infty} E[\mathbf{1}_A\langle X_t^{n_k,x}, \varphi'\rangle] = -\lim_{k\to\infty} E[\mathbf{1}_A\langle \frac{\partial}{\partial x} X_t^{n_k,x}, \varphi\rangle] = -E[\mathbf{1}_A\langle Y_t', \varphi\rangle].$$

So

$$\langle X_t, \varphi'\rangle = -\langle Y_t', \varphi\rangle, \quad P - a.s. \tag{6.83}$$

Finally, we need to show that there exists a measurable set $\Omega_0 \subset \Omega$ with full measure such that $X_t^\cdot$ has a weak derivative on this subset. To this end choose a sequence $\{\varphi_n\}$ in $C^\infty(\mathbb{R})$ dense in $W^{1,2}(U)$. Choose a measurable subset Ω_n of Ω with full measure such that (6.83) holds on Ω_n with φ replaced by φ_n. Then $\Omega_0 := \cap_{n\geq 1}\Omega_n$ satisfies the desired property.

Corollary 6.37. *Let $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ as in (6.6) and X_t^x the corresponding strong solution of* (6.5). *Then, for any compact subset $K \subset \mathbb{R}$ and $p \geq 1$*

$$\sup_{x\in K}\sup_{t\in[0,T]} E\left[\left(\frac{\partial}{\partial x}X_t^x\right)^p\right] \leq C_{K,p}$$

for a constant $C_{K,p} > 0$ depending on K and p. Here, $\frac{\partial}{\partial x}X_t^x$ is the first variation process of X_t^x, (see Proposition 6.18).

Proof. This is a direct consequence of Lemma 6.36 in connection with Fatou's lemma. $\square$

6.A.6 Proof of Proposition 6.18:

By Theorem 6.17 we know that the sequence $\{X_t^{n,x}\}_{n\geq 0}$ converges weakly to X_t^x in $L^2(\Omega, W^{1,2}(U))$. Therefore, it is enough to check that our candidate is the limit of $\frac{\partial}{\partial x}X_t^{n,x}$ in the weak topology of $L^2(U \times \Omega)$ for any open bounded $U \subset \mathbb{R}$, i.e.

$$\int_U \mathcal{W}\left(\exp\left\{ -\int_0^t \int_{\mathbb{R}} b_n(u,y) L^{X^{n,x}}(du,dy) \right\} - \exp\left\{ -\int_0^t \int_{\mathbb{R}} b(u,y) L^{X^x}(du,dy) \right\} \right)(f) g(x) dx$$

converges to 0 as $n \to \infty$ for every $f \in L^2([0,T])$ and $g \in C_0^\infty(U)$. This can be shown following exactly the same steps as in Proposition 6.15 by integrating I_n and II_n against $g(x)$ over $x \in U$. The only difference here is that we need all bounds to be uniformly in $x \in U$. At the end, one needs to show that

$$\sup_{n\geq 0} \sup_{x \in \mathrm{supp}(U)} E\left[|V_n|^{2p} e^{-2p\int_0^t \int_{\mathbb{R}} \tilde{b}_n(u,y) L^{B^x}(du,dy)} \right] < \infty$$

where

$$V_n := \int_0^t \int_{\mathbb{R}} (\tilde{b}_n(u,y) - \tilde{b}(u,y)) L^{B^x}(du,dy)$$

which holds by Lemma 6.31 and the fact that $\tilde{b}_n$, $n \geq 0$, is uniformly bounded. For II_n one can follow similar steps and use Remark 6.30.

Acknowledgement

We would like to thank Aaron Folly for support with the simulations in Section 6.5.

Chapter 7

Strong existence and higher order differentiability of stochastic flows of fractional Brownian motion driven SDEs with singular drift

David R. Baños, Torstein Nilssen and Frank Proske

Abstract: In this paper we present a new method for the construction of strong solutions of SDE's with merely integrable drift coefficients driven by a multidimensional fractional Brownian motion with Hurst parameter $H < \frac{1}{2}$. Furthermore, we prove the rather surprising result of the higher order Fréchet differentiability of stochastic flows of such SDE's in the case of a small Hurst parameter. In establishing these results we use techniques from Malliavin calculus combined with new ideas based on a "local time variational calculus". We expect that our general approach can be also applied to the study of certain types of stochastic partial differential equations as e.g. stochastic conservation laws driven by rough paths.

7.1 Introduction

Consider a d-dimensional fractional Brownian motion $B_t^H, t \geq 0$ with Hurst parameter $H \in (0,1)$ on a probability space $(\Omega, \mathcal{F}, P)$, that is a centered Gaussian process with a covariance structure $R_H(t,s)$ given by

$$R_H(t,s) = E[B_t^H B_s^H] = \frac{1}{2}\left(s^{2H} + t^{2H} - |t-s|^{2H}\right)$$

for all $t, s \geq 0$. The fractional Brownian motion, which is a Brownian motion in the case $H = \frac{1}{2}$, enjoys the property of self-similarity, that is

$$\{B_{\alpha t}^H\}_{t\geq 0} \overset{law}{=} \{\alpha^H B_t^H\}_{t\geq 0}$$

for all $\alpha > 0$. In fact the fractional Brownian motion, which has a version with $H^{-}-$continuous paths, is the only stationary Gaussian process satisfying the latter property. On the other hand this process is neither a Markov process nor a (weak) semimartingale and it is a very irregular process in the sense of rough paths for small Hurst parameters. See e.g. [90] and the references therein for more information about fractional Brownian motion.

In this article we aim at analysing solutions $X_{\cdot}^{x}$ of the stochastic differential equation (SDE)

$$X_t^x = x + \int_0^t b(X_s^x)ds + B_t^H, \quad 0 \le t \le T, \quad x \in \mathbb{R}^d, \tag{7.1}$$

where $B_{\cdot}^{H}$ is a fractional Brownian motion with Hurst parameter $H \in (0, \frac{1}{2})$ with respect to a P-augmented filtration $\{\mathcal{F}_t\}_{0\le t\le T}$ generated by $B_{\cdot}^{H}$ and where $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a Borel-measurable function.

If we impose a global Lipschitz and a linear growth condition on the drift coefficient b in (7.1), we can use the Picard iteration scheme to obtain a unique global strong solution to the SDE (7.1), that is a $\mathcal{F}_t-$adapted solution X_t^x to (7.1), which is a measurable $L^2(\Omega)$-functional of the driving noise.

However, a variety of important applications of such SDE's to stochastic control theory (in the case of $H = \frac{1}{2}$) (see [66]) or to the statistical mechanics of infinite particle systems (see [68]) show that the use of SDE's with regular coefficients in the sense of Lipschitzianity as models for random phenomena is not suitable and that one is forced to study such equations with coefficients which are irregular, that is discontinuous or merely measurable.

One objective of our paper is the construction of unique strong solutions to the SDE (7.1) driven by rough paths in the case of multidimensional fractional noise $B_{\cdot}^{H}$ for Hurst parameters $H < \frac{1}{2}$ and drift coefficients

$$b \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d).$$

In proving this new result, we employ tools from Malliavin Calculus and local time techniques.

The analysis of strong solutions to (7.1) has been a very active field of research in various branches of mathematics over the last decades. A foundational result in this direction of research was first obtained by Zvonkin in the beginning of the 1970ties [110], who showed the existence of a unique strong solution of one-dimensional Brownian motion driven SDE's (7.1), when the drift coefficient b is merely bounded and measurable. A few years later on the latter result was generalized by Veretennikov [105] to the multidimensional case.

More recently, Krylov and Röckner [68] gave the construction of unique strong solutions to (7.1) under integrability conditions on the (time-inhomogeneous) drift coefficient b. See also the articles [55] or [54]. In this context, we shall also mention the generalization of Zvonkin's result to the case of stochastic evolution equations in Hilbert spaces with bounded and measurable drift coefficients [28], where the authors use solutions to infinite-dimensional Kolmogorov equations to recast the singular drift term of the evolution equation in terms of a more regular expression ("Itô-Tanaka-Zvonkin trick").

In all of the above mentioned works the common technique of the authors for the construction of strong solutions rests on the so-called Yamada-Watanabe principle (see [108]), which entails strong uniqueness of solutions to SDE's, if pathwise uniqueness of (weak) solutions holds.

In fact, in order to ensure strong uniqueness of solutions, the above authors construct weak solutions to SDE's, which are not necessarily Brownian functionals, and verify pathwise uniqueness by using solutions of parabolic partial differential equations (see e.g. [110], [105] or [68]) or Skorohod embedding combined with Krylov's estimates (see e.g. [55], [54]).

We remark that the techniques of these authors for proving pathwise uniqueness are not applicable to SDE's driven by fractional Brownian motion, since the fractional Brownian is neither a Markov process nor a semimartingale for Hurst parameters $H \neq \frac{1}{2}$.

Further, we emphasize that our method, which is not only limited to Markov or semimartingale solutions of SDE's, gives a direct construction of strong solutions and provides a construction principle, which can be considered the converse to that of Yamada-Watanabe: We prove the existence of strong solutions and uniqueness in law to guarantee strong uniqueness.

To the best of our knowledge strong solutions to SDE's (7.1) for Hurst parameters $H < \frac{1}{2}$ and dimension $d \geq 2$ with irregular drift coefficients are for the first time obtained in this paper.

The case $d = 1$ for Hurst parameters $H \in (0,1)$ was treated in [91], where the authors prove strong uniqueness under a linear growth condition on the drift coefficient in the case $H < \frac{1}{2}$ by invoking a method based on the comparison theorem.

Another crucial objective of our article is the study of the regularity of stochastic flows of the SDE (7.1), that is the regularity of

$$(x \longmapsto X_t^x)$$

in the initial condition $x \in \mathbb{R}^d$, when the vector field b is discontinuous.

The motivation for this study comes from the deterministic case:

$$\frac{d}{dt} X_t^x = u(t, X_t^x), \quad X_0^x = x, \tag{7.2}$$

where $u : [0, \infty) \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is a vector field.

Here the solution $X : [0, \infty) \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ to (7.2) may e.g. stand for the flow of fluid particles with respect to the velocity field of an incompressible inviscid fluid whose dynamics is described by an incompressible Euler equation

$$u_t + (Du)u + \triangledown P = 0, \quad \nabla \cdot u = 0, \tag{7.3}$$

where $P : [0, \infty) \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is the pressure field.

Solutions to (7.3) may be singular. Therefore a better understanding of the regularity of solutions to (7.3) requires the study of flows of ODE's (7.2) driven by irregular vector fields.

If u is Lipschitz continuous it is well known that the unique flow $X : [0, \infty) \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ in (7.2) is Lipschitzian. The latter classical result was generalized by Di Perna and Lions in their celebrated paper [37] to the case of $u \in W^{1,p}$ and $\nabla \cdot u \in L^\infty$, for which the authors construct a unique flow X to (7.2). Later on the latter result was extended by Ambrosio [2] to the case of vector fields of bounded variation.

However, it turns out that the superposition of the ODE (7.2) by a Brownian noise $B_\cdot$, that is

$$dX_t = u(t, X_t)dt + dB_t, \quad s, t \in \mathbb{R}, \quad X_s = x \in \mathbb{R}^d \tag{7.4}$$

has a strong regularizing effect on its flow $\mathbb{R}^d \ni x \longmapsto \varphi_{s,t}(x) \in \mathbb{R}^d$.

Using techniques similar to those in this paper, but without arguments based on local time, it was shown in Mohammed, Nilssen, Proske [87] for merely *bounded measurable* drift coefficients u that $\varphi_{s,t}$ is a stochastic flow of Sobolev diffeomorphisms with

$$\varphi_{s,t}(\cdot), \varphi_{s,t}^{-1}(\cdot) \in L^2(\Omega, W^{1,p}(\mathbb{R}^d; w))$$

for all s, t and $p \in (1, \infty)$, where $W^{1,p}(\mathbb{R}^d; w)$ is a weighted Sobolev space with weight function $w : \mathbb{R}^d \longrightarrow [0, \infty)$.

As an application of this result the authors constructed Sobolev differentiable unique (weak) solutions to (Stratonovich) stochastic transport equation with multiplicative noise of the form

$$\begin{cases} d_t v(t,x) + (u(t,x) \cdot Dv(t,x))dt + \sum_{i=1}^d e_i \cdot Dv(t,x) \circ dB_t^i = 0 \\ u(0,x) = u_0(x), \end{cases}$$

where u is bounded and measurable, $u_0 \in C_b^1$ and where $\{e_i\}_{i=1}^d$ is a basis of $\mathbb{R}^d$.

By adopting ideas in Mohammed et al. [87], we mention that the latter result on the existence of stochastic flows of Sobolev diffeomorphisms was extended in [101] to the case of globally integrable $u \in L^{r,q}$ for $r/d+2/q < 1$ and applied to the study of the regularity of Navier-Stokes-equations. Compare also to [45], where the authors employ techniques based on solutions of backward Kolmogorov equations.

If the Brownian motion in (7.4) is replaced by a rougher noise given by $B_\cdot^H$ for $H < \frac{1}{2}$, we find in this paper for $u \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ the rather surprising result which generalises the classical result of Kunita [70] for smooth coefficients, that the stochastic flow $X : [0, \infty) \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is higher order Fréchet differentiable in the spatial variable, that is

$$(x \longmapsto X_t^x(\omega)) \in C^k(\mathbb{R}^d)$$

a.s. for all t and for $k \geq 1$, provided $H = H(k)$ is small enough.

In view of the above discussion in the case of Brownian noise driven stochastic flows, the latter result raises the fundamental question whether rough noise in the sense of $B_\cdot^H$ or a related noise with very irregular path behaviour may considerably regularise solutions of PDE's as e.g. transport equations, conservation laws or even Navier-Stokes equations by perturbation. We are confident that there is an affirmative answer for a class of interesting PDE's.

Finally, we comment on that the method for the construction of higher order Fréchet differentiable stochastic flows of (7.1), which is- as mentioned above- different from common techniques based on Markov processes and semimartingales, is inspired by the works [83], [81], [87], [56] in the case of (7.1) with initial Lévy noise and [48], [89] in the case of stochastic partial differential equations.

More precisely, in order to construct strong solutions to (7.1) we apply a compactness criterion for square integrable Brownian functionals in [30] to solutions X_t^n of

$$dX_t^n = b_n(X_t^n)dt + dB_t^H,$$

where $b_n, n \geq 1$ are smooth coefficients converging to b in $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ and show that X_t^n converges to a solution X_t of (7.1) in $L^2(\Omega)$ for all t.

In proving the latter and the higher order Fréchet differentiability of the corresponding stochastic flow we make use of a "local time variational calculus" argument of the form

$$\int_{\Delta_{\theta,t}^m} \kappa(s) D^\alpha f(B_s^H) ds = \int_{\mathbb{R}^{dm}} D^\alpha f(z) L_\kappa(t,z) dz = (-1)^{|\alpha|} \int_{\mathbb{R}^{dm}} f(z) D^\alpha L_\kappa(t,z) dz,$$

for smooth functions $f : \mathbb{R}^{dm} \longrightarrow \mathbb{R}$, where $L_\kappa(t,z)$ is a spatially differentiable local time on the simplex $\Delta_{\theta,t}^m = \{(s_1, ..., s_m) \in [0,T]^m : \theta < s_1 < ... < s_m < t\}$, scaled by a function κ (D^α is the partial derivative of order $|\alpha|$).

We expect that our approach can be also applied to the study of solutions of the following stochastic equations:

$$dX_t = (AX_t + b(X_t))dt + QdW_t^H,$$

for (mild) solutions X_t, where A is a densely defined linear operator (of parabolic type) on a separable Hilbert space H, $b : H \longrightarrow H$ is an irregular function, Q a Hilbert-Schmidt operator and $W_\cdot^H$ a cylindrical fractional Brownian motion.

On the other hand, using our method we may also examine equations of the type

$$dX_t = dA_t + dB_t^H,$$

where A_t is a process of bounded variation which arises from limits of the form

$$\lim_{n \to \infty} \int_0^t b_n(X_s) ds$$

for coefficients $b_n, n \geq 1$. See [16] in the Brownian case.

Our paper is organized as follows: In Section 2 we introduce the mathematical framework of the article and define in Section 3 the concept of a scaled local time of a fractional Brownian motion on a simplex, which we show to be high-order differentiable in the spatial variable for small Hurst parameters. In Section 4 we establish the existence of a unique strong solution to the SDE (7.1) under integrability conditions on the drift coefficient b. Section 5 is devoted to the study of the regularity properties of stochastic flows of (7.1).

7.2 Framework

In this section we recollect some specifics on fractional calculus, Malliavin calculus for fractional Brownian noise and occupation measures which will be extensively used throughout the article. The reader might consult [79], [78] or [36] for general theory on Malliavin calculus for Brownian motion and [90, Chapter 5] for fractional Brownian motion. Whereas for occupation measures one may review [52] or [62]. We present the results in one dimension for simplicity inasmuch as we will treat the multidimensional case.

7.2.1 Fractional calculus

We establish here some basic definitions and properties on fractional calculus. A general theory on this subject may be found in [102] and [86].

Let $a, b \in \mathbb{R}$ with $a < b$. Let $f \in L^p([a,b])$ wih $p \geq 1$ and $\alpha > 0$. Define the *left*- and *right-sided Riemann-Liouville fractional integrals* by

$$I_{a^+}^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x-y)^{\alpha-1} f(y) dy$$

and

$$I_{b^-}^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_x^b (y-x)^{\alpha-1} f(y) dy$$

for almost all $x \in [a,b]$ where Γ is the Gamma function.

Moreover, for a given integer $p \geq 1$, let $I_{a^+}^{\alpha}(L^p)$ (resp. $I_{b^-}^{\alpha}(L^p)$) denote the image of $L^p([a,b])$ by the operator $I_{a^+}^{\alpha}$ (resp. $I_{b^-}^{\alpha}$). If $f \in I_{a^+}^{\alpha}(L^p)$ (resp. $f \in I_{b^-}^{\alpha}(L^p)$) and $0 < \alpha < 1$ then define the *left*- and *right-sided Riemann-Liouville fractional derivatives* by

$$D_{a^+}^{\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \frac{\mathrm{d}}{\mathrm{d}x} \int_a^x \frac{f(y)}{(x-y)^{\alpha}} dy$$

and

$$D_{b^-}^{\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \frac{\mathrm{d}}{\mathrm{d}x} \int_x^b \frac{f(y)}{(y-x)^{\alpha}} dy.$$

The left- and right-sided derivatives of f defined above have the following representations

$$D_{a^+}^{\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(x-a)^{\alpha}} + \alpha \int_a^x \frac{f(x)-f(y)}{(x-y)^{\alpha+1}} dy \right)$$

and

$$D_{b^-}^{\alpha} f(x) = \frac{1}{\Gamma(1-\alpha)} \left(\frac{f(x)}{(b-x)^{\alpha}} + \alpha \int_x^b \frac{f(x)-f(y)}{(y-x)^{\alpha+1}} dy \right).$$

Finally, observe that by construction, the following formulas hold

$$I_{a^+}^{\alpha}(D_{a^+}^{\alpha} f) = f$$

for all $f \in I_{a^+}^{\alpha}(L^p)$ and

$$D_{a^+}^{\alpha}(I_{a^+}^{\alpha} f) = f$$

for all $f \in L^p([a,b])$ and similarly for $I_{b^-}^{\alpha}$ and $D_{b^-}^{\alpha}$.

7.2.2 Shuffles

Let $k \in \mathbb{N}$. For given $m_1, \ldots, m_k \in \mathbb{N}$, denote

$$m_j^- := \sum_{i=1}^{j} m_i,$$

e.g. $m_k^- = m_1 + \cdots + m_k$ and set $m_0 := 0$. Denote by $S_m = \{\sigma : \{1, \ldots, m\} \to \{1, \ldots, m\}\}$ the set of permutations of length $m \in \mathbb{N}$. Define the set of *shuffle permutations* of length $m_k^- = m_1 + \cdots m_k$ as

$$S(m_1, \ldots, m_k) := \{\sigma \in S_{m_k^-} : \sigma(m_i^- + 1) < \cdots < \sigma(m_{i+1}^-),\ i = 0, \ldots, k-1\}.$$

Fix $\theta, t \in [0, T]$ with $\theta < t$ and define the m-dimensional subset of $[0, T]^m$

$$\Delta_{\theta,t}^m := \{(s_1, \ldots, s_m) \in [0, T]^m : \theta < s_1 < \cdots < s_m < t\}.$$

Let $f_i : [0, T] \to [0, \infty)$, $i = 1, \ldots, m_k^-$ be integrable functions. Then, we have

$$\begin{aligned} &\prod_{i=0}^{k-1} \int_{\Delta_{\theta,t}^{m_i}} f_{m_i^-+1}(s_{m_i^-+1}) \cdots f_{m_{i+1}^-}(s_{m_{i+1}^-}) ds_{m_i^-+1} \cdots ds_{m_{i+1}^-} \\ &\qquad\qquad = \sum_{\sigma^{-1} \in S(m_1, \ldots, m_k)} \int_{\Delta_{\theta,t}^{m_k^-}} \prod_{i=1}^{m_k^-} f_{\sigma(i)}(w_i) dw_1 \cdots dw_{m_k^-}. \end{aligned} \tag{7.5}$$

The above is a trivial generalisation of the case $k = 2$ where

$$\begin{aligned} &\int_{\substack{\theta < s_1 \cdots < s_{m_1} < t \\ \theta < s_{m_1+1} < \cdots < s_{m_1+m_2} < t}} \prod_{i=1}^{m_1+m_2} f_i(s_i)\, ds_1 \cdots ds_{m_1+m_2} \\ &\qquad = \sum_{\sigma^{-1} \in S(m_1, m_2)} \int_{\theta < w_1 < \cdots < w_{m_1+m_2} < t} \prod_{i=1}^{m_1+m_2} f_{\sigma(i)}(w_i) dw_1 \cdots dw_{m_1+m_2} \end{aligned} \tag{7.6}$$

since

$$\begin{aligned} &\left\{(s_1, \ldots, s_{m_i}) \in [0, T]^{m_i}, i = 1, 2 : \theta < s_1 < \cdots < s_{m_i} < t, i = 1, 2\right\} \\ &\quad = \dot{\bigcup_{\sigma \in S(m_1, m_2)}} \left\{(w_1, \ldots, w_{m_1+m_2}) \in [0, T]^{m_1+m_2} : \theta < w_{\sigma(1)} < \cdots < w_{\sigma(m_1+m_2)} < t\right\}, \end{aligned}$$

which can also be found in [77, Theorem 2.15].

We will also need the following formula. Given indices $j_0, j_1, \ldots, j_{k-1} \in \mathbb{N}$ such that $1 \le j_i \le m_{i+1}$, $i = 1, \ldots, k-1$ and we set $j_0 := m_1 + 1$. Introduce the subset $S_{j_1, \ldots, j_{k-1}}(m_1, \ldots, m_k)$

of $S(m_1,\dots,m_k)$ defined as

$$S_{j_1,\dots,j_{k-1}}(m_1,\dots,m_k) := \Big\{\sigma\in S(m_1,\dots,m_k) : \sigma(m_i^-+1)<\cdots<\sigma(m_i^-+j_i-1), \\ \sigma(l)=l,\, m_i^-+j_i\le l\le m_{i+1}^-,\, i=0,\dots,k-1\Big\}.$$

We have

$$\begin{aligned}
&\int_{\Delta^{m_k}_{\theta,t}\times\Delta^{m_{k-1}}_{\theta,s_{m^-_{k-1}+j_{k-1}}}\times\cdots\times\Delta^{m_1}_{\theta,s_{m_1+j_1}}} \prod_{i=1}^{m_k^-} f_i(s_i)\,ds_1\cdots ds_{m_k^-} \\
&= \int_{\substack{\theta<s_1<\cdots<s_{m_1}<s_{m_1+j_1}\\ \theta<s_{m_1+m_2+1}<\cdots<s_{m_1+m_2}<s_{m_1+m_2+j_2}\\ \vdots \\ \theta<s_{m_1+\cdots m_{k-1}+1}<\cdots<s_{m_1+\cdots+m_k}<t}} \prod_{i=1}^{m_k^-} f_i(s_i)\,ds_1\cdots ds_{m_k^-} \\
&= \sum_{\sigma^{-1}\in S_{j_1,\dots,j_{k-1}}(m_1,\dots,m_k)} \int_{\theta<w_1<\cdots<w_{m_k^-}<t} \prod_{i=1}^{m_k^-} f_{\sigma(i)}(w_i)\,dw_1\cdots dw_{m_k^-}.
\end{aligned} \tag{7.7}$$

Observe that

$$\#S(m_1,\dots,m_k) = \frac{(m_1+\cdots+m_k)!}{m_1!\cdots m_k!}$$

where $\#$ denotes the number of elements in the given set. Then by using Stirling's approximation, one can show that

$$\#S(m_1,\dots,m_k)\le C^{m_1+\cdots+m_k}$$

for a large enough constant $C>0$. Moreover,

$$\#S_{j_1,\dots,j_{k-1}}(m_1,\dots,m_k)\le \#S(m_1,\dots,m_k).$$

7.2.3 Fractional Brownian motion

Let $B^H=\{B^H_t, t\in[0,T]\}$ be a d-dimensional *fractional Brownian motion* with Hurst parameter $H\in(0,1/2)$. In other words, B^H is a centered Gaussian process with covariance structure

$$R_H(t,s) := E[B^H_t B^H_s] = \frac{1}{2}\left(t^{2H}+s^{2H}-|t-s|^{2H}\right).$$

Observe that $E[|B^H_t-B^H_s|^2]=|t-s|^{2H}$ and hence B^H has stationary increments and Hölder continuous trajectories of index $H-\varepsilon$ for all $\varepsilon\in(0,H)$. Observe moreover that the increments of B^H, $H\in(0,1/2)$ are not independent. This fact makes computations more difficult due to the fact that B^H is not Markovian. Another difficulty one encounters is that B^H is not a semimartingale, see e.g. [90, Proposition 5.1.1].

Now we give a brief survey on how to construct fractional Brownian motion via an isometry. Since the construction can be done componentwise we present here for simplicity the one-dimensional case. Further details can be found in [90].

Denote by $\mathcal{E}$ the set of step functions on $[0,T]$ and denote by $\mathcal{H}$ the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the inner product

$$\langle 1_{[0,t]}, 1_{[0,s]}\rangle_{\mathcal{H}} = R_H(t,s).$$

The mapping $1_{[0,t]} \mapsto B_t$ can be extended to an isometry between $\mathcal{H}$ and the Gaussian subspace of $L^2(\Omega)$ associated with B^H. Denote such isometry by $\varphi \mapsto B^H(\varphi)$. We recall the following result (see [90, Proposition 5.1.3]) which gives an integral representation of $R_H(t,s)$ when $H < 1/2$.

Proposition 7.1. *Let $H < 1/2$. The kernel*

$$K_H(t,s) = c_H\left[\left(\frac{t}{s}\right)^{H-\frac{1}{2}}(t-s)^{H-\frac{1}{2}} - \left(H-\frac{1}{2}\right)s^{\frac{1}{2}-H}\int_s^t u^{H-\frac{3}{2}}(u-s)^{H-\frac{1}{2}}du\right]$$

where $c_H = \sqrt{\frac{2H}{(1-2H)\beta(1-2H,H+1/2)}}$ being β the Beta function, satisfies

$$R_H(t,s) = \int_0^{t\wedge s} K_H(t,u)K_H(s,u)du. \tag{7.8}$$

The kernel K_H can also be represented by means of fractional derivatives as follows

$$K_H(t,s) = c_H\Gamma\left(H+\frac{1}{2}\right)s^{\frac{1}{2}-H}\left(D_{t-}^{\frac{1}{2}-H}u^{H-\frac{1}{2}}\right)(s)$$

Consider the linear operator $K_H^* : \mathcal{E} \to L^2([0,T])$ defined by

$$(K_H^*\varphi)(s) = K_H(T,s)\varphi(s) + \int_s^T (\varphi(t)-\varphi(s))\frac{\partial K_H}{\partial t}(t,s)dt$$

for every $\varphi \in \mathcal{E}$. Observe that $(K_H^* 1_{[0,t]})(s) = K_H(t,s)1_{[0,t]}(s)$, then from this fact and (7.8) we see that K_H^* is an isometry between $\mathcal{E}$ and $L^2([0,T])$ which can be extended to the Hilbert space $\mathcal{H}$.

For a given $\varphi \in \mathcal{H}$ one can show the following two representations for K_H^* in terms of fractional derivatives

$$(K_H^*\varphi)(s) = c_H\Gamma\left(H+\frac{1}{2}\right)s^{\frac{1}{2}-H}\left(D_{T-}^{\frac{1}{2}-H}u^{H-\frac{1}{2}}\varphi(u)\right)(s)$$

and

$$\begin{aligned}(K_H^*\varphi)(s) =& c_H\Gamma\left(H+\frac{1}{2}\right)\left(D_{T-}^{\frac{1}{2}-H}\varphi(s)\right)(s)\\ &+ c_H\left(\frac{1}{2}-H\right)\int_s^T \varphi(t)(t-s)^{H-\frac{3}{2}}\left(1-\left(\frac{t}{s}\right)^{H-\frac{1}{2}}\right)dt.\end{aligned}$$

One can show that $\mathcal{H} = I_{T-}^{\frac{1}{2}-H}(L^2)$ (see [33] and [1, Proposition 6]).

Given the fact that K_H^* provides with an isometry from $\mathcal{H}$ into $L^2([0,T])$ the d-dimensional process $W = \{W_t, t \in [0,T]\}$ defined by

$$W_t := B^H((K_H^*)^{-1}(1_{[0,t]})) \tag{7.9}$$

is a Wiener process and the process B^H has the following representation

$$B_t^H = \int_0^t K_H(t,s)dW_s, \tag{7.10}$$

see [1].

Henceforward, we will denote by W a standard Wiener process on a given probability space $(\Omega, \mathcal{F}, P)$ equipped with the natural filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ generated by W augmented by all P-null sets and $B := B^H$ the fractional Brownian motion with Hurst parameter $H \in (0, 1/2)$ given by the representation (7.10).

Next, we give a version of Girsanov's theorem for fractional Brownian motion which is due to [33, Theorem 4.9]. Here we present the version given in [91, Theorem 3.1] but first we need to define an isomorphism K_H from $L^2([0,T])$ onto $I_{0+}^{H+\frac{1}{2}}(L^2)$ associated with the kernel $K_H(t,s)$ in terms of the fractional integrals as follows, see [33, Theorem 2.1]

$$(K_H\varphi)(s) = I_{0+}^{2H} s^{\frac{1}{2}-H} I_{0+}^{\frac{1}{2}-H} s^{H-\frac{1}{2}}\varphi, \quad \varphi \in L^2([0,T]).$$

From this and the properties of the Riemann-Liouville fractional integrals and derivatives the inverse of K_H is given by

$$(K_H^{-1}\varphi)(s) = s^{\frac{1}{2}-H} D_{0+}^{\frac{1}{2}-H} s^{H-\frac{1}{2}} D_{0+}^{2H}\varphi(s), \quad \varphi \in I_{0+}^{H+\frac{1}{2}}(L^2).$$

It follows that if φ is absolutely continuous, see [91], one can show that

$$(K_H^{-1}\varphi)(s) = s^{H-\frac{1}{2}} I_{0+}^{\frac{1}{2}-H} s^{\frac{1}{2}-H}\varphi'(s). \tag{7.11}$$

Theorem 7.2 (Girsanov's theorem for fBm)**.** *Let $u = \{u_t, t \in [0,T]\}$ be an $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted process with integrable trajectories and set $\widetilde{B}_t^H = B_t^H + \int_0^t u_s ds, \quad t \in [0,T]$. Assume that*

(i) $\int_0^\cdot u_s ds \in I_{0+}^{H+\frac{1}{2}}(L^2([0,T]))$, *$P$-a.s.*

(ii) $E[\xi_T] = 1$ *where*

$$\xi_T := \exp\left\{ -\int_0^T K_H^{-1}\left(\int_0^\cdot u_r dr\right)(s)dW_s - \frac{1}{2}\int_0^T K_H^{-1}\left(\int_0^\cdot u_r dr\right)^2(s)ds \right\}.$$

Then the shifted process $\widetilde{B}^H$ is an $\{\mathcal{F}_t\}_{t\in[0,T]}$-fractional Brownian motion with Hurst parameter H under the new probability $\widetilde{P}$ defined by $\frac{d\widetilde{P}}{dP} = \xi_T$.

Remark 7.3. *For the multidimensional case, define*

$$(K_H\varphi)(s) := ((K_H\varphi^{(1)})(s), \ldots, (K_H\varphi^{(d)})(s))^*, \quad \varphi \in L^2([0,T], \mathbb{R}^d),$$

where $*$ denotes transposition. Similarly for K_H^{-1} and K_H^*.

Finally, we will use a crucial property of the fractional Brownian motion which was proven by [97] for general Gaussian vector fields. This property will essentially help us to overcome the limitations of not having independent increments of the underlying noise.

Let $m \in \mathbb{N}$ and $0 =: t_0 < t_1 < \cdots < t_m < T$. Then for every $\xi_1, \ldots, \xi_m \in \mathbb{R}^d$ there exists a positive finite constant $C > 0$ (not depending on m) such that

$$Var\left[\sum_{j=1}^{m} \langle \xi_j, B_{t_j} - B_{t_{j-1}} \rangle_{\mathbb{R}^d}\right] \geq C \sum_{j=1}^{m} |\xi_j|^2 Var\left[|B_{t_j} - B_{t_{j-1}}|^2\right]. \tag{7.12}$$

The above property is known as the *(strong) local non-determinism* property of the fractional Brownian motion. The reader may consult [97] or [107] for more information on this property.

7.3 A local-time formula

From this point and on we will denote by $\mathcal{S}(\mathbb{R}^d; \mathbb{R}^d)$ (or simply $\mathcal{S}(\mathbb{R}^d)$ when the range is one-dimensional) the Schwarz class of functions, i.e. the space of rapidly decreasing functions

$$\mathcal{S}(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R}^d : \ f \in C^\infty(\mathbb{R}^d); \ \|f\|_{\alpha,\beta} < \infty, \forall \alpha, \beta\},$$

where

$$\|f\|_{\alpha,\beta} := \sup_{x \in \mathbb{R}^d} |x^\alpha D^\beta f(x)|$$

and where α and β are multi-indices and $x^\alpha := x_1^{\alpha_1} \cdots x_d^{\alpha_d}$, $x \in \mathbb{R}^d$, $\alpha = (\alpha_1, \ldots, \alpha_d)$. Finally, we denote by $C_c^\infty(\mathbb{R}^d)$ the subset of $\mathcal{S}(\mathbb{R}^d)$ of infinitely many times differentiable functions with compact support.

Let $m \in \mathbb{N}$ be fixed and $\kappa_j : [0,T] \to \mathbb{R}$, $j = 1, \ldots, m$ be measurable and positive functions. From now on, we fix $\theta, t \in [0,T]$ with $\theta < t$ and define the m-dimensional subset of $[0,T]^m$

$$\Delta_{\theta,t}^m := \{(s_1, \ldots, s_m) \in [0,T]^m : \ \theta < s_1 < \cdots < s_m < t\}.$$

We will henceforward use the following notation. Given $s = (s_1, \ldots, s_m) \in [0,T]^m$ and given a permutation $\sigma \in S_m$ we denote

$$\kappa(s) := \kappa_1(s_1) \cdots \kappa_m(s_m)$$

and

$$\kappa_\sigma(s) := \kappa_{\sigma(1)}(s_1) \cdots \kappa_{\sigma(m)}(s_m).$$

Further notations are used as in Section 7.2.2.

Let $\mathcal{B}(\mathbb{R}^{dm})$ be the Borel σ-algebra of $\mathbb{R}^{dm}$. Define the following occupation measure on $(\mathbb{R}^{dm}, \mathcal{B}(\mathbb{R}^{dm}))$ over $\Delta_{\theta,t}^m$ scaled by the function κ as follows:

$$\nu_\kappa^m(t, A) := \int_{\Delta_{\theta,t}^m} \kappa(s) \mathbf{1}_{\{B_s \in A\}} ds, \quad A \in \mathcal{B}(\mathbb{R}^{dm}) \tag{7.13}$$

where here B_s denotes the vector $B_s := (B_{s_1}, \ldots, B_{s_m})$.

Equivalently,

$$\int_{\Delta^m_{\theta,t}} \kappa(s) f(B_s) ds = \int_{\mathbb{R}^{dm}} f(x) \nu^m_\kappa(t, dx) \tag{7.14}$$

for all $f \in \mathcal{S}(\mathbb{R}^{dm})$.

The aim of this section is to show that there is a random field $L^m_\kappa : [0,T] \times \mathbb{R}^{dm} \times \Omega \to \mathbb{R}$ such that the following occupation-type formula holds

$$\int_{\Delta^m_{\theta,t}} \kappa(s) f(B_s) ds = \int_{\mathbb{R}^{dm}} f(z) L^m_\kappa(t, z) dz, \tag{7.15}$$

i.e. ν^m_κ is absolutely continuous w.r.t. Lebesgue measure.

We would like to establish a sufficient condition for the local time to exist and being several times differentiable w.r.t. the space variable. Indeed, let $k \geq 0$ integer and define

$$\Psi^m_{\kappa,k}(\theta, t) := \sum_{\sigma^{-1} \in S(m,m)} \int_{\Delta^{2m}_{\theta,t}} \kappa_\sigma(s) \prod_{j=1}^{2m} |s_j - s_{j-1}|^{-d(2k+1)H} ds_1 \cdots ds_{2m}, \tag{7.16}$$

where $s_0 := \theta$.

Introduce the following notation: $\alpha \in \mathbb{N}_0^{d \times m}$ and here $|\alpha| := \sum_{j=1}^m \sum_{l=1}^d \alpha_j^{(l)}$ and $D^\alpha L(t,z)$ denotes differentiation with respect to the space variable $z \in \mathbb{R}^{dm}$, i.e.

$$D^\alpha L^m_\kappa(t, z) = \prod_{j=1}^m \prod_{l=1}^d \frac{\partial^{\alpha_j^{(l)}}}{\partial (z_j^{(l)})^{\alpha_j^{(l)}}} L^m_\kappa(t, z), \quad z \in \mathbb{R}^{dm}. \tag{7.17}$$

It turns out that if $\Psi^m_{\kappa,k}(\theta, t) < \infty$ then the local time $L^m_\kappa(t, z)$ exists and is k-times differentiable w.r.t. $z \in \mathbb{R}^{dm}$. This is the content of the next theorem.

Theorem 7.4. *Let $m \in \mathbb{N}$ and $\kappa : [0,T]^m \to \mathbb{R}$ be a measurable positive function. Let $\alpha \in \mathbb{N}_0^{dm}$ such that $\alpha_j^{(l)} \leq k$ for every $j = 1, \ldots, m$ and $l = 1, \ldots, d$. Fix $\theta, t \in [0,T]$ and assume the following condition on κ is fulfilled*

$$\Psi^m_{\kappa,k}(\theta, t) := \sum_{\sigma^{-1} \in S(m,m)} \int_{\Delta^{2m}_{\theta,t}} \kappa_\sigma(s) \prod_{j=1}^{2m} |s_j - s_{j-1}|^{-d(2k+1)H} ds_1 \cdots ds_{2m} < \infty,$$

for some integer $k \geq 0$, where $s_0 := \theta$ and the sum is taken over $S(m,m)$, the set of all shuffle permutations of length $2m$. See Section 7.2.2.

Then there exists a k-times weakly differentiable function $z \mapsto L^m_\kappa(t, z)$ such that for every $f \in \mathcal{S}(\mathbb{R}^{dm})$ the following identity holds P-a.s.

$$\int_{\Delta^m_{\theta,t}} \kappa(s) f(B_s) ds = \int_{\mathbb{R}^{dm}} f(z) L^m_\kappa(t, z) dz.$$

Moreover, the following estimate holds true

$$\sup_{z\in\mathbb{R}^{dm}} E[|D^\alpha L^m_\kappa(t,z)|^2] \le C^m \Psi^m_{\kappa,k}(\theta,t). \tag{7.18}$$

We will refer to $L^m_\kappa(t,z)$, $t\in[0,T]$, $z\in\mathbb{R}^{dm}$ as the local-time of B^H over the simplex $\Delta^m_{\theta,t}$ at $z\in\mathbb{R}^{dm}$ scaled by κ.

Proof. Define

$$L^m_\kappa(t,z) := (2\pi)^{-dm}\int_{\mathbb{R}^{dm}}\int_{\Delta^m_{\theta,t}} \kappa(s)\exp\left\{-i\sum_{j=1}^m \langle v_j, B_{s_j}-z_j\rangle_{\mathbb{R}^d}\right\} dsdv. \tag{7.19}$$

Now we show that $L^m_\kappa(t,z)$ lies in $L^2(\Omega)$. Since the computations are the same for L^m_κ and its derivatives we will show the estimate (7.18) then one obtains $E[|L^m_\kappa(t,z)|^2]<\infty$ by choosing $k=0$.

Formal differentiation of $L^m_\kappa(t,z)$ yields

$$D^\alpha L^m_\kappa(t,z) = (2\pi)^{-dm}\int_{\Delta^m_{\theta,t}}\int_{(\mathbb{R}^d)^m} \kappa(s) i^{|\alpha|} u^\alpha e^{-i\sum_{j=1}^m \langle u_j, B_{s_j}-z_j\rangle_{\mathbb{R}^d}} duds.$$

Taking squared modulus

$$\begin{aligned}
&|D^\alpha L^m_\kappa(t,z)|^2 = \\
&=(2\pi)^{-2dm}\int_{(\mathbb{R}^d)^m}\int_{(\mathbb{R}^d)^m}\int_{\Delta^m_{\theta,t}} \kappa_1(s_1)\cdots\kappa_m(s_m)u^\alpha \prod_{j=1}^m e^{-i\langle u_j,B_{s_j}-z_j\rangle_{\mathbb{R}^d}} ds_1\cdots ds_m \\
&\qquad\times\int_{\Delta^m_{\theta,t}} \kappa_{m+1}(s_{m+1})\cdots\kappa_{2m}(s_{2m}) v^\alpha \prod_{j=m+1}^{2m} e^{i\langle v_j, B_{s_j}-z_j\rangle_{\mathbb{R}^d}} ds_{m+1}\cdots ds_{2m} dudv.
\end{aligned}$$

Use the following change of variables

$$\xi = \begin{pmatrix} I_{dm\times dm} & 0 \\ 0 & -I_{dm\times dm}\end{pmatrix}\begin{pmatrix} u \\ v\end{pmatrix}, \quad \xi\in(\mathbb{R}^d)^{2m}$$

and denote $\widetilde{\alpha} := (\alpha,\alpha)\in(\mathbb{N}_0^d)^{2m}$ to obtain that

$$\begin{aligned}
|D^\alpha L^m_\kappa(t,z)|^2 = &(-1)^{dm}(2\pi)^{-2dm}\int_{(\mathbb{R}^d)^{2m}} e^{i\sum_{j=1}^m \langle \xi_j+\xi_{j+m}, z_j\rangle_{\mathbb{R}^d}} \xi^{\widetilde{\alpha}} \\
&\times\left(\int_{\Delta^m_{\theta,t}} \kappa_1(s_1)\cdots\kappa_m(s_m)\prod_{j=1}^m f_j(s_j) ds_1\cdots ds_m\right) \\
&\times\left(\int_{\Delta^m_{\theta,t}} \kappa_{m+1}(s_{m+1})\cdots\kappa_{2m}(s_{2m})\prod_{j=m+1}^{2m} f_j(s_j) ds_{m+1}\cdots ds_{2m}\right) d\xi
\end{aligned}$$

where $f_j(s) = e^{-i\langle \xi_j, B_s\rangle_{\mathbb{R}^d}}$, $j=1,\dots,2m$.

Now shuffling the integrals w.r.t. s, see Section 7.2.2, and using the fact that f_j are symmetric we may write

$$
\begin{aligned}
&|D^\alpha L_\kappa^m(t,z)|^2 = \\
&= (-1)^{dm}(2\pi)^{-2dm}(-1)^{|\alpha|} \sum_{\sigma^{-1}\in S(m,m)} \int_{(\mathbb{R}^d)^{2m}} e^{i\sum_{j=1}^m \langle z_j, \xi_j+\xi_{j+m}\rangle_{\mathbb{R}^d}} \xi^{\widetilde{\alpha}} \\
&\times \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) \prod_{j=1}^{2m} e^{-i\langle \xi_j, B_{s_j}\rangle_{\mathbb{R}^d}} ds_1\cdots ds_{2m} d\xi.
\end{aligned}
$$

Taking expectation we then obtain

$$
\begin{aligned}
&E[|D^\alpha L_\kappa^m(t,z)|^2] \leq \\
&\leq (2\pi)^{-2dm} \sum_{\sigma^{-1}\in S(m,m)} \int_{(\mathbb{R}^d)^{2m}} |\xi^{\widetilde{\alpha}}| \left| \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) E\left[e^{-i\sum_{j=1}^{2m}\langle \xi_j, B_{s_j}\rangle_{\mathbb{R}^d}}\right] ds \right| d\xi
\end{aligned}
$$

Now apply the change of variables $\xi_{2m} = \zeta_{2m}$ and $\xi_j = \zeta_j - \zeta_{j+1}$, $j = 0, \ldots, 2m-1$, where we set $\zeta_0 := \zeta_1$, $\zeta_{2m+1} = 0$, $s_0 := \theta$ and $s_{-1} := 0$ in order to have

$$
\begin{aligned}
E[|D^\alpha L_\kappa^m(t,z)|^2] \leq & \sum_{\sigma^{-1}\in S(m,m)} (2\pi)^{-2dm} \int_{(\mathbb{R}^d)^{2m}} \prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)} - \zeta_{j+1}^{(l)}|^{\widetilde{\alpha}_j^{(l)}} \\
& \times \left| \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) E\left[e^{-i\sum_{j=0}^{2m}\langle \zeta_j, B_{s_j}-B_{s_{j-1}}\rangle_{\mathbb{R}^d}}\right] ds d\zeta \right| \\
= & \sum_{\sigma^{-1}\in S(m,m)} (2\pi)^{-2dm} \int_{(\mathbb{R}^d)^{2m}} \prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)} - \zeta_{j+1}^{(l)}|^{\widetilde{\alpha}_j^{(l)}} \\
& \times \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) e^{-\frac{1}{2}Var\left[\sum_{j=0}^{2m}\langle \zeta_j, B_{s_j}-B_{s_{j-1}}\rangle_{\mathbb{R}^d}\right]} ds d\zeta \\
\leq & \sum_{\sigma^{-1}\in S(m,m)} (2\pi)^{-2dm} \int_{(\mathbb{R}^d)^{2m}} \prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)} - \zeta_{j+1}^{(l)}|^{\widetilde{\alpha}_j^{(l)}} \\
& \times \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) e^{-\frac{C}{2}\sum_{j=0}^{2m} |\zeta_j|^2 |s_j - s_{j-1}|^{2H}} ds d\zeta
\end{aligned}
$$

for some constant $C > 0$ where we have used the strong local non-determinism of the fractional Brownian motion given in (7.12), see [97].

Now observe that we can express the product appearing in the integral above as a sum of different combinations where the exponent is, at most, two. That is

$$
\prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)} - \zeta_{j+1}^{(l)}|^{\widetilde{\alpha}_j^{(l)}} = \sum_{\substack{\delta\in I \\ \delta\in\{0,1,2\}^{2m}}} c_\delta \prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)}|^{\delta_j \widetilde{\alpha}_j^{(l)}}
$$

for some constants c_δ and here I is a set of indices which has 2^m elements. Now, we have that the integral w.r.t. ζ can be written as

$$\begin{aligned}A :=& \int_{(\mathbb{R}^d)^{2m}} \prod_{j=1}^{2m}\prod_{l=1}^{d} |\zeta_j^{(l)} - \zeta_{j+1}^{(l)}|^{\widetilde{\alpha}_j^{(l)}} e^{-\frac{C}{2}\sum_{j=0}^{2m}|\zeta_j|^2|s_j-s_{j-1}|^{2H}} d\zeta \\ =& \sum_{\substack{\delta\in I\\ \delta\in\{0,1,2\}^{2m}}} c_\delta \prod_{j=1}^{2m}\prod_{l=1}^{d}\left(\int_{\mathbb{R}} |\zeta_j^{(l)}|^{\delta_j\widetilde{\alpha}_j^{(l)}} e^{-\frac{1}{2\sigma_j^2}|\zeta_j^{(l)}|^2} d\zeta_j^{(l)}\right),\end{aligned}$$

where for each $j = 1,\dots,2m$,

$$\sigma_1^2 := \frac{1}{C}\left(\theta^{2H} + |s_1-\theta|^{2H}\right)^{-1} \le \frac{1}{C}|s_1-\theta|^{-2H}, \quad \sigma_j^2 := \frac{1}{C}|s_j - s_{j-1}|^{-2H}.$$

Then

$$\begin{aligned}A =& \sum_{\substack{\delta\in I\\ \delta\in\{0,1,2\}^{2m}}} c_\delta \prod_{j=1}^{2m}\prod_{l=1}^{d}\left(\sqrt{2\pi\sigma_j^2}\,\sigma_j^{\delta_j\widetilde{\alpha}_j^{(l)}} \frac{2^{\delta_j\widetilde{\alpha}_j^{(l)}/2}\Gamma\left(\frac{\delta_j\widetilde{\alpha}_j^{(l)}+1}{2}\right)}{\sqrt{\pi}}\right) \\ \le& C \sum_{\substack{\delta\in I\\ \delta\in\{0,1,2\}^{2m}}} \prod_{j=1}^{2m} \sigma_j^{\delta_j\sum_{l=1}^d \widetilde{\alpha}_j^{(l)}+d} \\ =& C^m \sum_{\substack{\delta\in I\\ \delta\in\{0,1,2\}^{2m}}} \prod_{j=1}^{2m}\prod_{l=1}^{d} |s_j - s_{j-1}|^{-H\left(\delta_j\sum_{l=1}^d \widetilde{\alpha}_j^{(l)}+d\right)} \\ \le& C^m \prod_{j=1}^{2m} |s_j - s_{j-1}|^{-dH(2k+1)} 1_{\{|s_j-s_{j-1}|<1\}} \\ &+ C^m 1_{\{|s_j-s_{j-1}|>1\}}\end{aligned}$$

for some constant $C>0$ depending only on d and k where we used $\sum_{l=1}^d \widetilde{\alpha}_j^{(l)} \le kd$ and $\delta_j \le 2$ for every $j = 1,\dots,2m$. The second term is clearly integrable w.r.t. s. Hence, we require that

$$\sum_{\sigma^{-1}\in S(m,m)} \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) \prod_{j=1}^{2m} |s_j - s_{j-1}|^{-dH(2k+1)} ds < \infty.$$

In particular, if $\kappa \equiv 1$ this is true if $H < \frac{1}{d(2k+1)}$.

As a matter of fact, we have

$$\begin{aligned}&\sup_{z\in\mathbb{R}^{dm}} E[|D^\alpha L_\kappa^m(t,z)|^2] \\ &\quad\le C^m \sum_{\sigma^{-1}\in S(m,m)} \int_{\Delta_{\theta,t}^{2m}} \kappa_\sigma(s) \prod_{j=1}^{2m} |s_j - s_{j-1}|^{-dH(2k+1)} ds.\end{aligned}$$

We remark that *a priori* one can not interchange the order of integration in (7.19). Indeed, for $m = 1$, $\kappa \equiv 1$ one gets an integral of the Donsker-Delta function. To overcome this define for $R > 0$

$$L^m_{\kappa,R}(t,z) := (2\pi)^{-dm} \int_{B(0,R)} \int_{\Delta^m_{\theta,t}} \kappa(s) e^{-i\sum_{j=1}^m \langle u_j, B_{s_j} - z_j\rangle_{\mathbb{R}^d}} ds dv$$

where $B(0,R) := \{v \in \mathbb{R}^{dm} : |v| < R\}$. Similar computations as above show that $L^m_{\kappa,R}(t,z) \to L^m_\kappa(t,z)$ in $L^2(\Omega)$ as $R \to \infty$ for all t and x.

Now for $f \in \mathcal{S}(\mathbb{R}^{dm})$, Lebesgue's dominated convergence theorem and the fact that the Fourier transform is an automorphism on the Schwarz space yield

$$\begin{aligned}
\int_{\mathbb{R}^{dm}} f(z) L^m_\kappa(t,z) dz &= \lim_{R\to\infty} \int_{\mathbb{R}^{dm}} f(z) L^m_{\kappa,R}(t,z) dx \\
&= \lim_{R\to\infty} (2\pi)^{-dm} \int_{\mathbb{R}^{dm}} \int_{B(0,R)} \int_{\Delta^m_{\theta,t}} f(z)\kappa(s) e^{-i\sum_{j=1}^m \langle u_j, B_{s_j} - z_j\rangle_{\mathbb{R}^d}} dzduds \\
&= \lim_{R\to\infty} \int_{\Delta^m_{\theta,t}} \kappa(s) \int_{B(0,R)} (2\pi)^{-dm} \int_{\mathbb{R}^{dm}} f(z) e^{-i\sum_{j=1}^m \langle u_j, B_{s_j}\rangle_{\mathbb{R}^d}} dzduds \\
&= \lim_{R\to\infty} \int_{\Delta^m_{\theta,t}} \kappa(s) \int_{B(0,R)} \widehat{f}(-u) e^{-i\sum_{j=1}^m \langle u_j, B_{s_j}\rangle_{\mathbb{R}^d}} duds \\
&= \int_{\Delta^m_{\theta,t}} \kappa(s) f(B_s) ds
\end{aligned}$$

which is exactly (7.15). □

Next, we give a crucial estimate which shows why fractional Brownian motion actually regularises (7.1). It is based on integration by parts and the aforementioned properties of the local-time L. The estimate we obtain can be presented in a more explicit way when

$$\kappa_j(s) = (K_H(s,\theta) - K_H(s,\theta'))^{\varepsilon_j}$$

for every $j = 1, \ldots, m$ or,

$$\kappa_j(s) = (K_H(s_j,\theta))^{\varepsilon_j}$$

for every $j = 1, \ldots, m$ with $(\varepsilon_1, \ldots, \varepsilon_m) \in \{0,1\}^m$ and we will see why these are important in the next coming section.

The proof can be found in the Appendix, Lemma 7.23 and Lemma 7.24.

Proposition 7.5. *Let B^H, $H \in (0, 1/2)$, be a standard d-dimensional fractional Brownian motion and $b_1, \ldots, b_m \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$. Let*

$$\kappa_j(s) = (K_H(s,\theta) - K_H(s,\theta'))^{\varepsilon_j}$$

for every $j = 1, \ldots, m$ with $(\varepsilon_1, \ldots, \varepsilon_m) \in \{0,1\}^m$ for $\theta, \theta' \in [0,T]$ with $\theta' < \theta$. Let $\alpha \in (\mathbb{N}_0^d)^m$ be an multi-index such that $\alpha_i^{(j)} \le k$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, d$. If

$$H < \min_{m\ge 1} \frac{m - \frac{1}{2}\sum_{j=1}^m \varepsilon_j}{md(2k+1) - \sum_{j=1}^m \varepsilon_j}, \tag{7.20}$$

then, there exists a universal constant $C > 0$ (independent of m, $\{b_i\}_{i=1,\dots,m}$ and α) such that for any $\theta, t \in [0,T]$ with $\theta < t$ we have

$$\begin{aligned}\Big| E\left[\int_{\Delta^m_{\theta,t}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B^H_{s_i})\kappa_i(s_i)\right) ds\right] \Big| \\ \leq C^m \prod_{i=1}^m \|b_i\|_{L^1(\mathbb{R}^d)} \frac{|\theta' - \theta|^{\gamma \sum_{j=1}^m \varepsilon_j} |t-\theta|^{m(1-d(2k+1)H) + \left(H - \frac{1}{2} - \gamma\right)\sum_{j=1}^m \varepsilon_j}}{\Gamma\left(2m(1-dH(2k+1)) + 1 + 2(H - \frac{1}{2} - \gamma)\sum_{j=1}^m \varepsilon_j\right)^{1/2}}\end{aligned} \tag{7.21}$$

for $\gamma \in (0, H)$.

Proof. By definition of L^m_κ in (7.15) it immediately follows that the integral in (7.21) can be expressed as

$$\int_{\Delta^m_{\theta,t}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B^H_{s_i})\kappa_i(s_i)\right) ds = \int_{\mathbb{R}^{dm}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(z_i)\right) L^m_\kappa(t,z)dz.$$

Now, since H satisfies (7.20) then it follows that (7.16) is finite and hence by Theorem 7.4 $z \mapsto L^m_\kappa(t,z)$, $z \in \mathbb{R}^{dm}$ is k-times differentiable. Because b_i, $i = 1, \dots, m$ are smooth with compact support we can use deterministic integration by parts to shift the derivatives on to the local-time $L^m_\kappa(\theta, \cdot)$, that is

$$\int_{\Delta^m_{\theta',\theta}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B^H_{s_i})\kappa_i(s_i)\right) ds = (-1)^{|\alpha|} \int_{\mathbb{R}^{dm}} \left(\prod_{i=1}^m b_i(z_i)\right) D^\alpha L^m_\kappa(t,z)dz.$$

Then taking expectation and absolute value we obtain

$$\left| E\left[\int_{\Delta^m_{\theta,t}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B^H_{s_i})\kappa_i(s_i)\right) ds\right]\right| \leq \int_{\mathbb{R}^{dm}} \left(\prod_{i=1}^m |b_i(z_i)|\right) E[|D^\alpha L^m_\kappa(t,z)|]dz. \tag{7.22}$$

Now, taking supremum we have

$$\left| E\left[\int_{\Delta^m_{\theta,t}} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B^H_{s_i})\kappa_i(s_i)\right) ds\right]\right| \leq \prod_{i=1}^m \|b_i\|_{L^1(\mathbb{R}^d)} \sup_{z\in\mathbb{R}^{dm}} E[|D^\alpha L^m_\kappa(t,z)|].$$

Finally, Theorem 7.4 and Lemma 7.23 in the Appendix allow us to conclude. □

Proposition 7.6. *Let B^H, $H \in (0, 1/2)$, be a standard d-dimensional fractional Brownian motion and $b_1, \dots, b_m \in C^\infty_c(\mathbb{R}^d, \mathbb{R})$ functions $b_i : \mathbb{R}^d \to \mathbb{R}$, $i = 1, \dots, m$ functions in the Schwarz space with compact support. Let $k : [0,T]^m \to \mathbb{R}$ be a chosen as*

$$\kappa_j(s) = (K_H(s,\theta))^{\varepsilon_j}$$

for every $j = 1, \dots, m$ with $(\varepsilon_1, \dots, \varepsilon_m) \in \{0,1\}^m$ for $\theta, \theta' \in [0,T]$ with $\theta' < \theta$. Let $\alpha \in$

$(\mathbb{N}_0^d)^m$ *be an multi-index such that* $\alpha_i^{(j)} \leq k$ *for all* $i = 1, \ldots, m$ *and* $j = 1, \ldots, d$. *If*

$$H < \min_{m \geq 1} \frac{m - \frac{1}{2}\sum_{j=1}^m \varepsilon_j}{md(2k+1) - \sum_{j=1}^m \varepsilon_j}, \tag{7.23}$$

then, there exists a universal constant $C > 0$ *(independent of* m, $\{b_i\}_{i=1,\ldots,m}$ *and* α*) such that for any* $\theta, t \in [0,T]$ *with* $\theta < t$ *we have*

$$\begin{aligned} &\left| E\left[\int_{\Delta_{\theta,t}^m} \left(\prod_{i=1}^m D^{\alpha_i} b_i(B_{s_i}^H)\kappa_i(s_i) \right) ds \right] \right| \\ &\leq C^m \prod_{i=1}^m \|b_i\|_{L^1(\mathbb{R}^d)} \frac{|t-\theta|^{m(1-dH(2k+1)) + \left(H - \frac{1}{2}\right)\sum_{j=1}^m \varepsilon_j}}{\Gamma\left(2m(1-dH(2k+1)) + 1 + 2(H-\frac{1}{2})\sum_{j=1}^m \varepsilon_j\right)^{1/2}} \end{aligned} \tag{7.24}$$

for $\gamma \in (0, H)$.

Proof. Similar to the proof of Proposition 7.5 in connection with Lemma 7.24 in the Appendix. □

7.4 Existence and uniqueness of global strong solutions

As outlined in the introduction the object of study is a time-homogeneous SDE with additive d-dimensional fractional Brownian noise B^H with Hurst parameter $H \in (0, 1/2)$, i.e.

$$X_t = x + \int_0^t b(X_s)ds + B_t^H, \quad t \in [0,T] \tag{7.25}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel-measurable function such that (7.25) makes sense, that is,

$$\int_0^T b(X_s)ds < \infty, \quad P-a.s. \tag{7.26}$$

We will study equation (7.1) when the drift coefficient b belongs to $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.

Definition 7.7. *Let* $x \in \mathbb{R}^d$ *and* $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *be a Borel-measurable function such that* (7.26) *holds. We say that a stochastic process* $X = \{X_t, t \in [0,T]\}$ *is a strong solution of* (7.25) *if*

$$X_t = x + \int_0^t b(X_s)ds + B_t^H$$

for every $t \in [0,T]$ *and* X *is adapted to* $\{\mathcal{F}_t\}_{t\in[0,T]}$ *the* P*-augmented filtration generated by* B^H.

Hereunder, we establish the main result of this section.

Theorem 7.8. *Let* $b \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. *Then if* $H < \frac{1}{2(3d-1)}$, $d \geq 1$ *there exists a unique (global) strong solution* $X = \{X_t, t \in [0,T]\}$ *of equation* (7.1). *Moreover, for every* $t \in [0,T]$, X_t *is Malliavin differentiable in the direction of the Brownian motion* W *in* (7.9).

The proof of Theorem 7.8 is based on the following steps:

1. First, we construct a weak solution X to (7.1) by means of Girsanov's theorem, that is we introduce a probability space $(\Omega, \mathcal{F}, P)$ that carries a fractional Brownian motion B^H and a process X such that (7.1) is fulfilled. However, a priori X is not adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ generated by B^H, which is the same as the filtration generated by W.

2. Next, we approximate the drift coefficient b by a sequence of functions (which always exists by standard approximation results) b_n, $n \geq 1$ such that $\{b_n\}_{n\geq 1} \subset C_c^\infty(\mathbb{R}^d)$ with $\|b_n - b\|_{L^1(\mathbb{R}^d)} \to 0$ as $n \to \infty$. By standard results on SDEs, we know that for each smooth coefficient b_n, $n \geq 1$, there exists unique strong solution $X_\cdot^n$ to the SDE

$$dX_t^n = b_n(X_t^n)du + dB_t^H,\ 0 \leq t \leq T,\ \ X_0^n = x \in \mathbb{R}^d\,. \tag{7.27}$$

   We then show that for each $t \in [0,T]$ the sequence X_t^n converges weakly to the conditional expectation $E[X_t|\mathcal{F}_t]$ in the space $L^2(\Omega;\mathcal{F}_t)$ of square integrable, $\mathcal{F}_t$-measurable random variables.

3. It is well known, see e.g. [90], that that for each $t \in [0,T]$ the strong solution X_t^n, $n \geq 1$, is Malliavin differentiable, and that the Malliavin derivative $D_s X_t^n$, $0 \leq s \leq t$, with respect to W in (7.9) satisfies

$$D_s X_t^n = K_H(t,s)I_d + \int_s^t b_n'(X_u^n)D_s X_u^n du, \tag{7.28}$$

   where b_n' denotes the Jacobian of b_n. In the next step we then employ a compactness criterion based on Malliavin calculus to show that for every $t \in [0,T]$ the set of random variables $\{X_t^n\}_{n\geq 0}$ is relatively compact in $L^2(\Omega;\mathcal{F}_t)$, which then admits the conclusion that X_t^n converges strongly in $L^2(\Omega;\mathcal{F}_t)$ to $E[X_t|\mathcal{F}_t]$. Further we see that $E[X_t|\mathcal{F}_t]$ is Malliavin differentiable as a consequence of the compactness criterion.

4. In the last step we show that $E[X_t|\mathcal{F}_t] = X_t$, which implies that X_t is $\mathcal{F}_t$-measurable and thus a strong solution.

We turn to the first step of our scheme which is to construct weak solutions of (7.1) by using Girsanov's theorem in this context. Let $(\Omega, \mathcal{F}, \widetilde{P})$ be some given probability space which carries a d-dimensional fractional Brownian motion $\widetilde{B}^H$ with Hurst parameter $H \in (0,1/2)$ and set $X_t := x + \widetilde{B}_t^H$, $t \in [0,T]$, $x \in \mathbb{R}^d$. Set $\theta_t := \left(K_H^{-1}\left(\int_0^\cdot b(X_r)dr\right)\right)(t)$ and consider the Doléans-Dade exponential

$$Z_t := \mathcal{E}\left(\theta_\cdot\right)_t := \exp\left\{\int_0^t \theta_s^T dW_s - \frac{1}{2}\int_0^t \theta_s^T\theta_s ds\right\}, \quad t \in [0,T].$$

The following two lemmata show that the conditions of Theorem 7.2 hold.

Lemma 7.9. *Let $\tilde{B}_t^H$ be a d-dimensional fractional Brownian motion with respect to $(\Omega, \mathcal{F}, \tilde{P})$. Then*

$$\int_0^\cdot |b(\tilde{B}_s^H)|ds \in I_{0+}^{H+\frac{1}{2}}(L^2), \quad P-a.s.$$

Proof. Using the property that $D_{0^+}^{H+\frac{1}{2}} I_{0^+}^{H+\frac{1}{2}}(f) = f$ for $f \in L^2([0,T])$ we need to show that

$$D_{0^+}^{H+\frac{1}{2}} \int_0^{\cdot} |b(\tilde{B}_s^H)| ds \in L^2([0,T]), \quad P-a.s.$$

Indeed,

$$\begin{aligned}\left| D_{0^+}^{H+\frac{1}{2}} \left(\int_0^{\cdot} |b(\tilde{B}_s^H)| ds \right)(t) \right| =& \frac{1}{\Gamma\left(\frac{1}{2}-H\right)} \left(\frac{1}{t^{H+\frac{1}{2}}} \int_0^t |b(\widetilde{B}_u^H)| du \right. \\ &\left. + \left(H+\frac{1}{2}\right) \int_0^t (t-s)^{-H-\frac{3}{2}} \int_s^t |b(\widetilde{B}_u^H)| ds \right) \\ =& \frac{1}{\Gamma\left(\frac{1}{2}-H\right)} \|b\|_\infty \left(t^{\frac{1}{2}-H} + \frac{H+\frac{1}{2}}{\frac{1}{2}-H} t^{\frac{1}{2}-H} \right).\end{aligned}$$

Hence, for some finite constant $C_H > 0$ we have

$$\left| D_{0^+}^{H+\frac{1}{2}} \left(\int_0^{\cdot} |b(\tilde{B}_s^H)| ds \right)(t) \right|^2 \leq C_H \|b\|_\infty^2 t^{1-2H}.$$

As a result,

$$\int_0^T \left| D_{0^+}^{H+\frac{1}{2}} \left(\int_0^{\cdot} |b(\tilde{B}_s^H)| ds \right)(t) \right|^2 dt \leq C_H \|b\|_\infty^2 \int_0^T t^{1-2H} dt < \infty, \quad P-a.s.$$

since $H \in (0, 1/2)$. $\square$

Lemma 7.10. *Let $\tilde{B}_t^H$ be a d-dimensional fractional Brownian motion with respect to $(\Omega, \mathcal{F}, \tilde{P})$. Then for every $\mu \geq 0$ we have*

$$E\left[\exp\left\{ \mu \int_0^T \left| K_H^{-1} \left(\int_0^{\cdot} b(\tilde{B}_r^H) dr \right)(s) \right|^2 ds \right\} \right] \leq C_{H,d,\mu,T}(\|b\|_{L^\infty(\mathbb{R}^d)})$$

for some continuous increasing function $C_{H,d,\mu,T}$ depending only on H, d, T and μ.

In particular,

$$E\left[\mathcal{E}\left(\int_0^T K_H^{-1} \left(\int_0^{\cdot} b(\tilde{B}_r^H) dr \right)^* (s) dW_s \right)^p \right] \leq C_{H,d,\mu,T}(\|b\|_{L^\infty(\mathbb{R}^d)})$$

where $*$ denotes transposition.

Proof. Denote by $\theta_s := K_H^{-1}\left(\int_0^{\cdot} |b(\tilde{B}_r^H)| dr \right)(s)$. Then using relation (7.11) we have

$$\begin{aligned}|\theta_s| =& |s^{H-\frac{1}{2}} I_{0^+}^{\frac{1}{2}-H} s^{\frac{1}{2}-H} |b(\widetilde{B}_s^H)|| \\ =& \frac{1}{\Gamma\left(\frac{1}{2}-H\right)} s^{H-\frac{1}{2}} \int_0^s (s-r)^{-\frac{1}{2}-H} r^{\frac{1}{2}-H} |b(\widetilde{B}_r^H)| dr \\ \leq& \|b\|_\infty \frac{1}{\Gamma\left(\frac{1}{2}-H\right)} s^{H-\frac{1}{2}} \int_0^s (s-r)^{-\frac{1}{2}-H} r^{\frac{1}{2}-H} dr\end{aligned}$$

$$
\begin{aligned}
&= \|b\|_\infty \frac{\Gamma\left(\frac{3}{2}-H\right)}{\Gamma(1-2H)} s^{\frac{1}{2}-H} \\
&\leq \|b\|_\infty \frac{\Gamma\left(\frac{3}{2}-H\right)}{\Gamma(1-2H)} T^{\frac{1}{2}-H}.
\end{aligned}
$$

Squaring both sides we have the following estimate

$$
|\theta_s|^2 \leq C_H \|b\|_\infty^2 T^{1-2H} \quad P-a.s. \tag{7.29}
$$

where $C_H := \frac{\Gamma\left(\frac{3}{2}-H\right)^2}{\Gamma(1-2H)^2}$.

Then using Taylor's expansion for the exponential function and the above estimate we have

$$
\begin{aligned}
&E\left[\exp\left\{\mu \int_0^T |\theta_s|^2 \, ds\right\}\right] = E\left[\sum_{m\geq 1} \mu^m \int_{\Delta_{0,T}^m} \prod_{i=1}^m |\theta_{s_i}|^2 ds_1 \cdots ds_m\right] \\
&\leq \sum_{m\geq 1} \mu^m \int_{\Delta_{0,T}^m} \left(\|b\|_\infty^2 C_H T^{1-2H}\right)^m ds_1 \cdots ds_m \\
&= \sum_{m\geq 1} \frac{\left(\mu C_H T^{2(1-H)} \|b\|_\infty^2\right)^m}{m!} \\
&= \exp\left\{\mu C_H T^{2(1-H)} \|b\|_\infty^2\right\}.
\end{aligned}
$$

□

By Girsanov's theorem, see Theorem 7.2, the process

$$
B_t^H := X_t - x - \int_0^t b(X_s) ds, \quad t \in [0,T] \tag{7.30}
$$

is a fractional Brownian motion on $(\Omega, \mathcal{F}, P)$ with Hurst parameter $H \in (0, 1/2)$. Hence, because of (7.30), the couple (X, B^H) is a weak solution of (7.1) on $(\Omega, \mathcal{F}, P)$.

Henceforth, we confine ourselves to the filtered probability space $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_t\}_{t\in[0,T]}$ which carries the weak solution (X, B^H) of (7.1).

Remark 7.11. *As outlined in the scheme above, the main challenge to establish existence of a strong solution is now to show that X is $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted. Indeed, in that case $X_t = F_t(B_\cdot)$ for some family of measurable functionals F_t, $t \in [0,T]$, (see e.g. [82] for an explicit form of F_t), and for any other stochastic basis $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \hat{B})$ one gets that $X_t := F_t(\hat{B}_\cdot)$, $t \in [0,T]$, is a $\hat{B}$-adapted solution to SDE (7.1). But this means exactly the existence of a strong solution to SDE (7.1).*

Remark 7.12. *It is worth to remark that one actually has existence of weak solutions for any $H \in (0, 1/2)$ and that weak solutions for bounded b are weakly unique since the estimates from Lemma 7.10 also hold with X in place of $\widetilde{B}^H$. For this reason, the main challenge is to show that when H is small enough such solutions are in fact strong. Then weak uniqueness implies strong uniqueness. See [100].*

We turn now to the second step of our procedure.

Lemma 7.13. *Let* $\{b_n\}_{n\geq 1} \subset C_c^\infty(\mathbb{R}^d)$ *be such that* $\lim_{n\to\infty} \|b_n - b\|_{L^1(\mathbb{R}^d)} = 0$. *Denote by* $X^n = \{X_t^n, t \in [0,T]\}$ *the corresponding solutions of* (7.1) *if we replace* b *by* b_n, $n \geq 1$. *Then for every* $t \in [0,T]$ *and globally Lipschitz continuous function* $\varphi : \mathbb{R}^d \to \mathbb{R}$ *we have that*

$$\varphi(X_t^n) \xrightarrow{n\to\infty} E\left[\varphi(X_t)|\mathcal{F}_t\right]$$

weakly in $L^2(\Omega; \mathcal{F}_t)$.

Proof. Let us first show that

$$\mathcal{E}\left(\int_0^t K_H^{-1}\left(\int_0^\cdot b_n(B_r^H)dr\right)^*(s)dW_s\right) \to \mathcal{E}\left(\int_0^t K_H^{-1}\left(\int_0^\cdot b(B_r^H)dr\right)^*(s)dW_s\right) \tag{7.31}$$

in $L^p(\Omega)$ for all $p \geq 1$. To see this, note that

$$K_H^{-1}\left(\int_0^\cdot b_n(B_r^H)dr\right)(s) \to K_H^{-1}\left(\int_0^\cdot b(B_r^H)dr\right)(s)$$

in probability for all s. Indeed, similar computations as in Lemma 7.10 give

$$\begin{aligned}
&E\left[\left|K_H^{-1}\left(\int_0^\cdot b_n(B_r^H)dr\right)(s) - K_H^{-1}\left(\int_0^\cdot b(B_r^H)dr\right)(s)\right|\right]\\
&\leq \frac{s^{H-1/2}}{\Gamma(\frac{1}{2}-H)}\int_0^s (s-r)^{-1/2-H} r^{1/2-H} E[|b_n(B_r^H) - b(B_r^H)|]dr\\
&= \frac{s^{H-1/2}}{\Gamma(\frac{1}{2}-H)}\int_0^s (s-r)^{-1/2-H} r^{1/2-H}\int_{\mathbb{R}^d} |b_n(y) - b(y)|(2\pi r^{2H})^{-d/2}\exp\left\{-\frac{y^2}{2r^{2H}}\right\}dydr\\
&\leq C_d\frac{s^{H-1/2}}{\Gamma(\frac{1}{2}-H)}\int_0^s (s-r)^{-1/2-H} r^{1/2-H(d+1)}dr\|b_n - b\|_{L^1(\mathbb{R}^d)} \to 0
\end{aligned}$$

as $n \to \infty$ since the above integral is finite when $H < \frac{3}{2(d+1)}$.

Moreover, $\left\{K_H^{-1}(\int_0^\cdot b_n(B_r^H)dr)\right\}_{n\geq 0}$ is bounded in $L^2([0,t]\times\Omega;\mathbb{R}^d)$. This is directly seen from (7.29) in Lemma 7.10.

Consequently

$$\int_0^t K_H^{-1}\left(\int_0^\cdot b_n(B_r^H)dr\right)^*(s)dW_s \to \int_0^t K_H^{-1}\left(\int_0^\cdot b(B_r^H)dr\right)^*(s)dW_s$$

and

$$\int_0^t \left|K_H^{-1}\left(\int_0^\cdot b_n(B_r^H)dr\right)(s)\right|^2 ds \to \int_0^t \left|K_H^{-1}\left(\int_0^\cdot b(B_r^H)dr\right)(s)\right|^2 ds$$

in $L^2(\Omega)$ since the latter is bounded $L^p(\Omega)$ for any $p \geq 1$, see Lemma 7.10.

Using the estimate $|e^x - e^y| \leq e^{x+y}|x-y|$, Hölder's inequality and the bounds in Lemma 7.10 it is clear that (7.31) holds.

Similarly, one also shows that

$$\exp\left\{\left\langle \alpha, \int_s^t b_n(B_r^H)dr\right\rangle\right\} \to \exp\left\{\left\langle \alpha, \int_s^t b(B_r^H)dr\right\rangle\right\}$$

in $L^p(\Omega)$ for all $p \geq 1, 0 \leq s \leq t \leq T, \alpha \in \mathbb{R}^d$.

To conclude the proof we note that the set

$$\Sigma_t := \left\{ \exp\{\sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H \rangle\} : \{\alpha_j\}_{j=1}^k \subset \mathbb{R}^d, 0 = t_0 < \cdots < t_k = t, k \geq 1 \right\}$$

is a total subspace of $L^2(\Omega, \mathcal{F}_t, P)$ and we may thus restrict ourselves to show the convergence

$$\lim_{n\to\infty} E\left[(\varphi(X_t^n) - E[\varphi(X_t)|\mathcal{F}_t])\, \xi \right] = 0$$

for all $\xi \in \Sigma_t$. To this end, we notice that φ is of linear growth and hence $\varphi(B_t)$ has all moments. Consequently we have the following convergence

$$E\left[\varphi(X_t^n) \exp\left\{ \sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H \rangle \right\} \right]$$

$$= E\left[\varphi(X_t^n) \exp\left\{ \sum_{j=1}^{k} \langle \alpha_j, X_{t_j}^n - X_{t_{j-1}}^n - \int_{t_{j-1}}^{t_j} b_n(X_s^n) ds \rangle \right\} \right]$$

$$= E[\varphi(B_t) \exp\{\sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H - \int_{t_{j-1}}^{t_j} b_n(B_s^H) ds \rangle\} \mathcal{E}\left(\int_0^t K_H^{-1}\left(\int_0^{\cdot} b_n(B_r^H) dr \right)(s) dW_s \right)]$$

$$\to E[\varphi(B_t) \exp\{\sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H - \int_{t_{j-1}}^{t_j} b(B_s^H) ds \rangle\} \mathcal{E}\left(\int_0^t K_H^{-1}\left(\int_0^{\cdot} b(B_r^H) dr \right)(s) dW_s \right)]$$

$$= E[\varphi(X_t) \exp\{\sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H \rangle\}]$$

$$= E[E[\varphi(X_t)|\mathcal{F}_t] \exp\{\sum_{j=1}^{k} \langle \alpha_j, B_{t_j}^H - B_{t_{j-1}}^H \rangle\}].$$

□

We continue to proving the third step of our scheme. This is the most challenging part. The following result is based on a compactness criterion for subsets of $L^2(\Omega)$ which is summarised in the Appendix.

Lemma 7.14. *Let $\{b_n\}_{n\geq 1} \subset C_c^\infty(\mathbb{R}^d)$ the approximating sequence of b in $L^1(\mathbb{R}^d)$. Denote by X_t^n the corresponding solutions of* (7.1) *if we replace b by b_n, $n \geq 1$. Fix $t \in [0, T]$ then there exists a $\beta \in (0, 1/2)$ such that*

$$\sup_{n\geq 1} \int_0^t \int_0^t \frac{E[\|D_\theta X_t^n - D_{\theta'} X_t^n\|^2]}{|\theta' - \theta|^{1+2\beta}} d\theta' d\theta < \infty$$

and

$$\sup_{n\geq 1}\|D.X_t^n\|_{L^2(\Omega\times[0,T])}<\infty. \tag{7.32}$$

Proof. Fix $t\in[0,T]$ and take $\theta,\theta'>0$ such that $0<\theta'<\theta<t$. Using the chain rule for the Malliavin derivative, see [90, Proposition 1.2.3], we have

$$D_\theta X_t^n = K_H(t,s)I_d+\int_\theta^t b_n'(X_s^n)D_\theta X_s^n ds$$

P-a.s. for all $0\leq\theta\leq t$ where $b_n'(z)=\left(\frac{\partial}{\partial z_j}b_n^{(i)}(z)\right)_{i,j=1,\dots,d}$ denotes the Jacobian matrix of b and I_d the identity matrix in $\mathbb{R}^{d\times d}$. Thus we have

$$\begin{aligned}
D_{\theta'}X_t^n - D_\theta X_t^n =& K_H(t,\theta')I_d - K_H(t,\theta)I_d\\
&+\int_{\theta'}^t b_n'(X_s^n)D_{\theta'}X_s^n ds - \int_\theta^t b_n'(X_s^n)D_\theta X_s^n ds\\
=&K_H(t,\theta')I_d - K_H(t,\theta)I_d\\
&+\int_{\theta'}^\theta b_n'(X_s^n)D_{\theta'}X_s^n ds + \int_\theta^t b_n'(X_s^n)(D_{\theta'}X_s^n - D_\theta X_s^n)ds\\
=&K_H(t,\theta')I_d - K_H(t,\theta)I_d + D_{\theta'}X_\theta^n - K_H(\theta,\theta')I_d\\
&+\int_\theta^t b_n'(X_s^n)(D_{\theta'}X_s^n - D_\theta X_s^n)ds.
\end{aligned}$$

Using Picard iteration applied to the above equation we may write

$$\begin{aligned}
D_{\theta'}X_t^n - D_\theta X_t^n =& K_H(t,\theta')I_d - K_H(t,\theta)I_d\\
&+\sum_{m=1}^\infty \int_{\Delta_{\theta,t}^m}\prod_{j=1}^m b_n'(X_{s_j}^n)\left(K_H(s_m,\theta')I_d - K_H(s_m,\theta)I_d\right)ds_m\cdots ds_1\\
&+\left(I_d+\sum_{m=1}^\infty \int_{\Delta_{\theta,t}^m}\prod_{j=1}^m b_n'(X_{s_j}^n)ds_m\cdots ds_1\right)\left(D_{\theta'}X_\theta^n - K_H(\theta,\theta')I_d\right).
\end{aligned}$$

On the other hand, observe that one may again write

$$D_{\theta'}X_\theta^n - K_H(\theta,\theta')I_d = \sum_{m=1}^\infty \int_{\Delta_{\theta',\theta}^m}\prod_{j=1}^m b_n'(X_{s_j}^n)(K_H(s_m,\theta')I_d)\,ds_m\cdots ds_1.$$

Altogether, we can write

$$\begin{aligned}
D_{\theta'}X_t^n - D_\theta X_t^n =& K_H(t,\theta')I_d - K_H(t,\theta)I_d \\
&+ \sum_{m=1}^{\infty} \int_{\Delta_{\theta,t}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n)\left(K_H(s_m,\theta')I_d - K_H(s_m,\theta)I_d\right) ds_m \cdots ds_1 \\
&+ \left(I_d + \sum_{m=1}^{\infty} \int_{\Delta_{\theta,t}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n) ds_m \cdots ds_1\right) \\
&\times \left(\sum_{m=1}^{\infty} \int_{\Delta_{\theta',\theta}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n)(K_H(s_m,\theta')I_d) ds_m \cdots ds_1.\right).
\end{aligned}$$

Introduce the notation $D_{\theta'}X_t^n - D_\theta X_t^n = I_1(\theta',\theta) + I_2^n(\theta',\theta) + I_3^n(\theta',\theta)$, where

$$\begin{aligned}
I_1(\theta',\theta) :=& K_H(t,\theta')I_d - K_H(t,\theta)I_d \\
I_2^n(\theta',\theta) :=& \sum_{m=1}^{\infty} \int_{\Delta_{\theta,t}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n)\left(K_H(s_m,\theta')I_d - K_H(s_m,\theta)I_d\right) ds_m \cdots ds_1 \\
I_3^n(\theta',\theta) :=& \left(I_d + \sum_{m=1}^{\infty} \int_{\Delta_{\theta,t}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n) ds_m \cdots ds_1\right) \\
&\times \left(\sum_{m=1}^{\infty} \int_{\Delta_{\theta',\theta}^m} \prod_{j=1}^{m} b_n'(X_{s_j}^n)(K_H(s_m,\theta')I_d) ds_m \cdots ds_1.\right).
\end{aligned}$$

It follows from Lemma 7.22 that

$$\int_0^t \int_0^t \frac{\|I_1(\theta',\theta)\|_{L^2(\Omega)}^2}{|\theta'-\theta|^{1+2\beta}} d\theta d\theta' = \int_0^t \int_0^t \frac{|K_H(t,\theta') - K_H(t,\theta)|^2}{|\theta'-\theta|^{1+2\beta}} d\theta d\theta' < \infty$$

for $\beta \in (0,1/2)$.

Let us continue with term $I_2^n(\theta',\theta)$. Then Girsanov's theorem, Cauchy-Schwarz inequality and Lemma 7.10 imply

$$\begin{aligned}
&E[\|I_2^n(\theta',\theta)\|^2] \\
&\quad \leq CE\left[\left\|\sum_{m=1}^{\infty} \int_{\Delta_{\theta,t}^m} \prod_{j=1}^{m} b_n'(x + B_{s_j}^H)\left(K_H(s_m,\theta')I_d - K_H(s_m,\theta)I_d\right) ds_m \cdots ds_1\right\|^4\right]^{1/2}.
\end{aligned}$$

for a finite constant $C > 0$.

Let $\|\cdot\|$ denote the matrix norm in $\mathbb{R}^{d\times d}$ such that $\|A\| = \sum_{i,j=1}^d |a_{ij}|$ for a matrix $A = \{a_{ij}\}_{i,j=1,\ldots,d}$, then taking this matrix norm and expectation we have

$$E[\|I_2^n(\theta',\theta)\|^2] \leq \left(\sum_{m=1}^{\infty} \sum_{i,j=1}^{d} \sum_{l_1,\dots,l_{m-1}=1}^{d} \left\| \int_{\Delta_{\theta,t}^m} \frac{\partial}{\partial x_{l_1}} b_n^{(i)}(x+B_{s_1}^H) \frac{\partial}{\partial x_{l_2}} b_n^{(l_1)}(x+B_{s_2}^H) \cdots \right.\right.$$
$$\left.\left. \cdots \frac{\partial}{\partial x_j} b_n^{(l_{m-1})}(x+B_{s_m}^H) \left(K_H(s_m,\theta') - K_H(s_m,\theta)\right) ds_m \cdots ds_1 \right\|_{L^4(\Omega,\mathbb{R})} \right)^2 .$$

Now look at the expression

$$J_2^n(\theta',\theta) := \int_{\Delta_{\theta,t}^m} \frac{\partial}{\partial x_{l_1}} b_n^{(i)}(x+B_{s_1}^H) \cdots \frac{\partial}{\partial x_j} b_n^{(l_{m-1})}(x+B_{s_m}^H) \left(K_H(s_m,\theta') - K_H(s_m,\theta)\right) ds. \tag{7.33}$$

Then, shuffling $J_2^n(\theta',\theta)$ as shown in (7.5), one can write $(J_2^n(\theta',\theta))^2$ as a sum of at most 2^{2m} summands of length $2m$ of the form

$$\int_{\Delta_{\theta,t}^{2m}} g_1^n(B_{s_1}^H) \cdots g_{2m}^n(B_{s_{2m}}^H) ds_{2m} \cdots ds_1, \tag{7.34}$$

where for each $l = 1,\dots,2m$,

$$g_l^n(B_\cdot^H) \in \left\{ \frac{\partial}{\partial x_j} b_n^{(i)}(x+B_\cdot^H), \frac{\partial}{\partial x_j} b_n^{(i)}(x+B_\cdot^H) \left(K_H(\cdot,\theta') - K_H(\cdot,\theta)\right), \, i,j = 1,\dots,d \right\}.$$

Repeating this argument once again, we find that $J_2^n(\theta',\theta)^4$ can be expressed as a sum of, at most, 2^{8m} summands of length $4m$ of the form

$$\int_{\Delta_{\theta,t}^{4m}} g_1^n(B_{s_1}^H) \cdots g_{4m}^n(B_{s_{4m}}^H) ds_{4m} \cdots ds_1, \tag{7.35}$$

where for each $l = 1,\dots,4m$,

$$g_l^n(B_\cdot^H) \in \left\{ \frac{\partial}{\partial x_j} b_n^{(i)}(x+B_\cdot^H), \frac{\partial}{\partial x_j} b_n^{(i)}(x+B_\cdot^H) \left(K_H(\cdot,\theta') - K_H(\cdot,\theta)\right), \, i,j = 1,\dots,d \right\}.$$

It is important to note that the function $(K_H(\cdot,\theta') - K_H(\cdot,\theta))$ appears only once in term (7.33) and hence only four times in term (7.35). So there are indices $j_1,\dots,j_4 \in \{1,\dots,4m\}$ such that we can write (7.35) as

$$\int_{\Delta_{\theta,t}^{4m}} \left(\prod_{j=1}^{4m} b_j^n(B_{s_j}^H) \right) \prod_{i=1}^{4} \left(K_H(s_{j_i},\theta') - K_H(s_{j_i},\theta)\right) ds_{4m} \cdots ds_1$$

where

$$b_l^n(B_\cdot^H) \in \left\{ \frac{\partial}{\partial x_j} b_n^{(i)}(x+B_\cdot^H), \, i,j = 1,\dots,d \right\}, \quad l = 1,\dots,4m.$$

The latter enables us to use the estimate from Proposition 7.5 with $\sum_{j=1}^{4m} \varepsilon_j = 4$ and $k = 1$

and thus we obtain that

$$E(J_2^n(\theta',\theta))^4 \leq 2^{8m}C^m\|b\|_{L^1(\mathbb{R}^d)}^{4m}\frac{|\theta'-\theta|^{4\gamma}|t-\theta|^{4m(1-3dH)+4\left(H-\frac{1}{2}-\gamma\right)}}{\Gamma\left(8m(1-3dH)+1+8(H-\frac{1}{2}-\gamma)\right)^{1/2}},$$

whenever $H<\frac{1}{6d-2}$.

Altogether, we see that

$$E\left[\|I_2^n(\theta',\theta)\|^2\right] \leq \left(\sum_{m=1}^{\infty} d^{m+1}2^{2m}C^m\frac{\|b_n\|_{L^1(\mathbb{R}^d)}^m|\theta'-\theta|^{\gamma}}{\Gamma\left(8m(1-3dH)+1+8(H-\frac{1}{2}-\gamma)\right)^{1/8}}\right)^2.$$

So we can find a constant $C>0$ such that

$$\sup_{n\geq 0} E\left[\|I_2^n(\theta',\theta)\|^2\right] \leq C|\theta'-\theta|^{\varepsilon}$$

for a small enough $\varepsilon\in(0,1)$ provided that $H<\frac{1}{2(3d-1)}$.

We turn now to term $I_3^n(\theta',\theta)$. Observe that term $I_3^n(\theta',\theta)$ is the product of two terms, where the first one will simply be bounded uniformly in $\theta,t\in[0,T]$ under expectation. This can be shown by following meticulously the same steps as we did for $I_2^n(\theta',\theta)$.

Again Girsanov's theorem, Cauchy-Schwarz inequality several times and Lemma 7.10 lead to

$$\begin{aligned}E[\|I_3^n(\theta',\theta)\|^2] \leq & C_2\left\|I_d+\sum_{m=1}^{\infty}\int_{\Delta_{\theta,t}^m}\prod_{j=1}^m b_n'(x+B_{s_j}^H)ds_m\cdots ds_1\right\|_{L^8(\Omega,\mathbb{R}^{d\times d})}^2\\ &\times\left\|\sum_{m=1}^{\infty}\int_{\Delta_{\theta',\theta}^m}\prod_{j=1}^m b_n'(x+B_{s_j}^H)K_H(s_m,\theta')ds_m\cdots ds_1\right\|_{L^4(\Omega,\mathbb{R}^{d\times d})}^2\end{aligned}$$

Again, we have

$$\begin{aligned}E[\|I_3^n(\theta',\theta)\|^2] \leq & C_2\Bigg(1+\sum_{m=1}^{\infty}\sum_{i,j=1}^{d}\sum_{l_1,\dots,l_{m-1}=1}^{d}\Bigg\|\int_{\Delta_{\theta,t}^m}\frac{\partial}{\partial x_{l_1}}b^{(i)}(x+B_{s_1}^H)\cdots\\ &\cdots\frac{\partial}{\partial x_j}b^{(l_{m-1})}(x+B_{s_m}^H)ds_m\cdots ds_1\Bigg\|_{L^8(\Omega,\mathbb{R})}\Bigg)^2\\ &\times\Bigg(\sum_{m=1}^{\infty}\sum_{i,j=1}^{d}\sum_{l_1,\dots,l_{m-1}=1}^{d}\Bigg\|\int_{\Delta_{\theta',\theta}^m}\frac{\partial}{\partial x_{l_1}}b^{(i)}(x+B_{s_1}^H)\cdots\\ &\cdots\frac{\partial}{\partial x_j}b^{(l_{m-1})}(x+B_{s_m}^H)K_H(s_m,\theta')ds_m\cdots ds_1\Bigg\|_{L^4(\Omega,\mathbb{R})}\Bigg)^2.\end{aligned}$$

Using exactly the same ideas as for $I_2^n(\theta',\theta)$ we see that the first factor can be bounded by some finite constant $C_3 > 0$, i.e.

$$E[\|I_3(\theta',\theta)\|^2] \leq C_3\Bigg(\sum_{m=1}^{\infty}\sum_{i,j=1}^{d}\sum_{l_1,\dots,l_{m-1}=1}^{d}\Bigg\|\int_{\Delta^m_{\theta',\theta}}\frac{\partial}{\partial x_{l_1}}b^{(i)}(x+B^H_{s_1})\cdots$$
$$\cdots\frac{\partial}{\partial x_j}b^{(l_{m-1})}(x+B^H_{s_m})K_H(s_m,\theta')ds_m\cdots ds_1\Bigg\|_{L^4(\Omega,\mathbb{R})}\Bigg)^2.$$

As before, look at

$$J_3^n(\theta',\theta) := \int_{\Delta^m_{\theta',\theta}}\frac{\partial}{\partial x_{l_1}}b_n^{(i)}(x+B^H_{s_1})\cdots\frac{\partial}{\partial x_j}b_n^{(l_{m-1})}(x+B^H_{s_m})K_H(s_m,\theta')ds_m\cdots ds_1. \quad (7.36)$$

We can express $(J_3(\theta',\theta))^4$ as a sum of, at most, 2^{8m} summands of length $4m$ of the form

$$\int_{\Delta^{4m}_{\theta',\theta}} g_1^n(B^H_{s_1})\cdots g_{4m}^n(B^H_{s_{4m}})ds_{4m}\cdots ds_1 \quad (7.37)$$

where for each $l = 1,\dots,4m$,

$$g_l^n(B^H_\cdot) \in \left\{\frac{\partial}{\partial x_j}b_n^{(i)}(x+B^H_\cdot),\, \frac{\partial}{\partial x_j}b_n^{(i)}(x+B^H_\cdot)K_H(\cdot,\theta'),\, i,j=1,\dots,d\right\},$$

where the factor $K_H(\cdot,\theta')$ is repeated four times in the integrand of (7.37). Now we can simply apply Proposition 7.6 with $k = 1$ and $\sum_{j=1}^{4m}\varepsilon_j = 4$ in order to get

$$E[(J_3^n(\theta',\theta))^8] \leq 2^{8m}C^m\|b\|^{4m}_{L^1(\mathbb{R}^d)}\frac{|\theta-\theta'|^{4m(1-3dH)+4\left(H-\frac{1}{2}\right)}}{\Gamma\left(8m(1-3dH)+1+8(H-\frac{1}{2})\right)^{1/2}}.$$

As a result,

$$E[\|I_3^n(\theta',\theta)\|^2] \leq \left(\sum_{m=1}^{\infty}d^{m+1}2^{2m}C^m\|b\|^m_{L^1(\mathbb{R}^d)}\frac{|\theta-\theta'|^{m(1-3dH)+H-\frac{1}{2}}}{\Gamma\left(8m(1-3dH)+1+4(H-\frac{1}{2})\right)^{1/8}}\right)^2.$$

Hence,

$$\sup_{n\geq 0}E[\|I_3^n(\theta',\theta)\|^2] \leq C|\theta'-\theta|^{\varepsilon}$$

for some $\varepsilon \in (0,1)$ small enough provided $H < \frac{1}{2(3d-1)}$.

Altogether,

$$\sup_{n\geq 0}\int_0^t\int_0^t\frac{E[\|D_{\theta'}X_t^n - D_\theta X_t^n\|^2]}{|\theta'-\theta|^{1+2\beta}}d\theta' d\theta < \infty$$

for $\beta \in (0,1/2)$.

Similar computations show that

$$\sup_{n\geq 0}\|D_\cdot X_t^n\|_{L^2(\Omega\times[0,T])} < \infty.$$

□

Corollary 7.15. *Let* $\{b_n\}_{n\geq 1} \subset C_c^\infty(\mathbb{R}^d)$ *the approximating sequence of* b *a.e. in* $L^1(\mathbb{R}^d)$. *Denote by* X_t^n *the corresponding solutions of* (7.1) *if we replace* b *by* b_n, $n \geq 0$. *Then for every* $t \in [0,T]$ *and globally Lipschitz continuous function* $\varphi : \mathbb{R}^d \to \mathbb{R}$ *we have*

$$\varphi(X_t^n) \xrightarrow{n\to\infty} \varphi(E\left[X_t|\mathcal{F}_t\right])$$

strongly in $L^2(\Omega;\mathcal{F}_t)$. *In addition,* $E\left[X_t|\mathcal{F}_t\right]$ *is Malliavin differentiable for every* $t \in [0,T]$.

Proof. This is an immediate consequence of the relatively compactness obtained in Lemma 7.14 and by Lemma 7.13 we can identify the limit as being $E[X_t|\mathcal{F}_t]$ then the convergence holds for any globally Lipschitz continuous functions as well. The Malliavin differentiability of $E[X_t|\mathcal{F}_t]$ is shown by taking $\varphi = I_d$ and estimate (7.32) together with [90, Proposition 1.2.3]. □

Finally, we can prove the main result of this section.

Proof of Theorem 7.8. It remains to prove that X_t is $\mathcal{F}_t$-measurable for every $t \in [0,T]$ and by Remark 7.11 it then follows that there exists a strong solution in the usual sense that is Malliavin differentiable. Indeed, let φ be a globally Lipschitz continuous function, then by Corollary 7.15 we have, for a subsequence n_k, $k \geq 0$, that

$$\varphi(X_t^{n_k}) \to \varphi(E[X_t|\mathcal{F}_t]), \quad P-a.s.$$

as $k \to \infty$.

On the other hand, by Lemma 7.13 we also have

$$\varphi(X_t^n) \to E\left[\varphi(X_t)|\mathcal{F}_t\right]$$

weakly in $L^2(\Omega;\mathcal{F}_t)$. By the uniqueness of the limit we immediately have

$$\varphi\left(E[X_t|\mathcal{F}_t]\right) = E\left[\varphi(X_t)|\mathcal{F}_t\right], \quad P-a.s.$$

which implies that X_t is $\mathcal{F}_t$-measurable for every $t \in [0,T]$.

Finally, to show uniqueness it is enough to show that two given strong solutions are weakly unique, indeed, one can follow the same argument as in [100, Chapter IX, Exercise (1.20)] which asserts that strong existence and uniqueness in law implies pathwise uniqueness. The argument does not rely on the process being a semimartingale. Since our solutions are, by construction, strong and uniqueness in law follows from Novikov's condition from Lemma 7.10 replacing B^H by X then pathwise uniqueness follows. □

7.5 Stochastic flows and regularity properties

Henceforward, we will denote $X_t^{s,x}$ the solution to the following SDE driven by fractional Brownian motion with $H < 1/2$

$$dX_t^{s,x} = b(X_t^{s,x})dt + dB_t^H, \quad s,t \in [0,T], \quad s \le t, \quad X_s^{s,x} = x \in \mathbb{R}^d. \tag{7.38}$$

We will then assume the hypotheses from Theorem 7.8 on b and H. The next result tells us that if $H = H(k)$ is small enough we may gain regularity on $x \mapsto X_t^{s,x}$. In particular, it shows that the strong solution constructed in the former section, in addition to Malliavin differentiability, is also once differentiable with respect to x since $k = 1$.

Theorem 7.16. *Let $b \in C_c^\infty(\mathbb{R}^d)$. Fix integers $p \ge 2$ and $k \ge 1$. Then, if $H < \frac{1}{d(2k+1)}$ we have*

$$\sup_{s,t\in[0,T]} \sup_{x\in\mathbb{R}^d} E\left[\left\| \frac{\partial^k}{\partial x^k} X_t^{s,x} \right\|^p\right] \le C(\|b\|_{L^1(\mathbb{R}^d)}),$$

where $C : [0,\infty) \to [0,\infty)$ is a continuous function, depending on k, d, H, p and T.

Proof. Given a k-times differentiable vector field $f : \mathbb{R}^d \to \mathbb{R}^d$, $f(x) = (f^{(1)}(x), \dots, f^{(d)}(x))$, we will indistinctively denote by

$$D^k f(x) = \frac{\partial^k}{\partial x^k} f(x) = \left\{ \frac{\partial^k}{\partial x_j \partial x_{l_1} \cdots \partial x_{l_{k-1}}} f^{(i)}(x) \right\}_{i,l_1,\dots,l_{k-1},j=1,\dots,d}$$

and identify $\mathbb{R}^d \times \overset{k+1)}{\cdots} \times \mathbb{R}^d \cong \mathbb{R}^{kd\times d}$.

Since $b \in C_c^\infty(\mathbb{R}^d)$, we know that the solution of (7.38), $X_t^{s,x}$ is smooth in the initial value x and that

$$\frac{\partial}{\partial x} X_t^{s,x} = I_d + \int_s^t Db(X_u^{s,x}) \frac{\partial}{\partial x} X_u^{s,x} du.$$

Using Picard's iteration we get

$$\frac{\partial}{\partial x} X_t^{s,x} = I_d + \sum_{m\ge1} \int_{\Delta_{s,t}^m} Db(X_{u_1}^{s,x}) \cdots Db(X_{u_m}^{s,x}) du_m \cdots du_1. \tag{7.39}$$

Now apply $\frac{\partial}{\partial x}$ again, then by dominated convergence we have

$$\frac{\partial^2}{\partial x^2} X_t^{s,x} = \sum_{m\ge1} \int_{\Delta_{s,t}^m} \frac{\partial}{\partial x} \left[Db(X_{u_1}^{s,x}) \cdots Db(X_{u_m}^{s,x}) \right] du_m \cdots du_1. \tag{7.40}$$

We can expand (7.40) using Leibniz's rule as follows

$$\frac{\partial}{\partial x}\left[Db(X_{u_1}^{s,x}) \cdots Db(X_{u_m}^{s,x}) \right] = \sum_{r=1}^m Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) \frac{\partial}{\partial x} X_{u_r}^{s,x} \cdots Db(X_{u_m}^{s,x}).$$

Inserting the representation (7.39) for $DX_t^{s,x}$ in this case we have that

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2} X_t^{s,x} &= \sum_{m_1 \geq 1} \int_{\Delta_{s,t}^{m_1}} \sum_{r=1}^{m_1} Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) \\
&\quad \times \left(I_d + \sum_{m_2 \geq 1} \int_{\Delta_{s,t}^{m_2}} Db(X_{v_1}^{s,x}) \cdots Db(X_{v_{m_2}}^{s,x}) dv_{m_2} \cdots dv_1 \right) \\
&\quad \times Db(X_{u_{r+1}}^{s,x}) \cdots Db(X_{u_{m_1}}^{s,x}) du_{m_1} \cdots du_1. \\
&= \sum_{m_1 \geq 1} \Bigg\{ \int_{\Delta_{s,t}^{m_1}} \sum_{r=1}^{m_1} Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) Db(X_{u_{r+1}}^{s,x}) \cdots Db(X_{u_{m_1}}^{s,x}) du_{m_1} \cdots du_1 \\
&\quad + \int_{\Delta_{s,t}^{m_1}} \sum_{r=1}^{m_1} Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) \\
&\quad \times \left(\sum_{m_2 \geq 1} \int_{\Delta_{s,u_r}^{m_2}} Db(X_{v_1}^{s,x}) \cdots Db(X_{v_{m_2}}^{s,x}) dv_{m_2} \cdots dv_1 \right) \\
&\quad \times Db(X_{u_{r+1}}^{s,x}) \cdots Db(X_{u_{m_1}}^{s,x}) du_{m_1} \cdots du_1 \Bigg\}.
\end{aligned}
$$

We reallocate terms by dominated convergence and respecting the order of matrices

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2} X_t^{s,x} &= \sum_{m_1 \geq 1} \sum_{r=1}^{m_1} \int_{\Delta_{s,t}^{m_1}} Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) \cdots Db(X_{u_{m_1}}^{s,x}) du_{m_1} \cdots du_1 \qquad (7.41) \\
&\quad + \sum_{m_1 \geq 1} \sum_{r=1}^{m_1} \sum_{m_2 \geq 1} \int_{\Delta_{s,t}^{m_1}} \int_{\Delta_{s,u_r}^{m_2}} Db(X_{u_1}^{s,x}) \cdots D^2 b(X_{u_r}^{s,x}) \\
&\quad \times Db(X_{v_1}^{s,x}) \cdots Db(X_{v_{m_2}}^{s,x}) Db(X_{u_{r+1}}^{s,x}) \cdots Db(X_{u_{m_1}}^{s,x}) dv_{m_2} \cdots dv_1 du_{m_1} \cdots du_1. \\
&=: I_1 + I_2
\end{aligned}
$$

where I_1 and I_2 denote the respective summands in the expression.

Now, we iterate this scheme, up to step $k \geq 2$. We will obtain that $\frac{\partial^k}{\partial x^k} X_t^{s,x}$ is a sum of 2^{k-1} terms. That is

$$
\frac{\partial^k}{\partial x^k} X_t^{s,x} = I_1 + \cdots + I_{2^{k-1}},
$$

where each I_i, $i = 1, \ldots, 2^{k-1}$ is an integral over at most $\Delta_{s,t}^{m_1 + \cdots + m_k}$ with at most one factor $D^k b$ and the rest $D^j b$, $j \leq k-1$.

In order to simplify the reading clearer we introduce some notation. For given indexes $m := (m_1, \ldots, m_k)$ and $r := (r_1, \ldots, r_{k-1})$ denote

$$
m_j^- := \sum_{i=1}^{j} m_i \quad \text{and} \quad m_j^+ := \sum_{i=j}^{k} m_i
$$

and

$$\sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}} := \sum_{m_1\geq 1}\sum_{r_1=1}^{m_1}\sum_{m_2\geq 1}\sum_{r_2=1}^{m_2^-}\cdots\sum_{r_{k-1}=1}^{m_{k-1}^-}\sum_{m_k\geq 1}.$$

Moreover, denote

$$\int_{\Delta^m}\cdot\, du := \int_{\Delta^{m_1}_{s,t}}\int_{\Delta^{m_2}_{s,u^1_{r_1}}}\cdots\int_{\Delta^{m_k}_{s,u^{k-1}_{r_{k-1}}}}\cdot\, du$$

where

$$u = (u^k_{m_k},\dots,u^k_1,\dots,u^1_{m_1},\dots,u^1_1)\in[0,T]^{m_1^+}.$$

Finally, given $\sigma := (\sigma_1,\dots,\sigma_{k-1})$ shuffle permutations. Denote

$$\sum_{\sigma^{-1}} := \sum_{\sigma^{-1}\in S_{r_1,\dots,r_{k-1}}(m_1,\dots,m_k)}. \tag{7.42}$$

If $f_i:[0,T]\to\mathbb{R}$, $i=1,\dots,m_1^+$ are measurable functions, then notice that, by using the argument in (7.7), one can express the integral over Δ^m as an expression of sums like the one given in (7.42) of integrals on simplices. That is, setting $m^+_{k+1}:=0$ one can write

$$\int_{\Delta^m}\prod_{j=1}^{k}\prod_{i=m^+_{j+1}+1}^{m_j^+} f_i(u_i^j)du = \sum_{\sigma^{-1}}\int_{\Delta^{m_1^+}_{s,t}}\prod_{i=1}^{m_1^+} f_{\sigma(i)}(w_i)dw. \tag{7.43}$$

Then $\frac{\partial^k}{\partial x^k}X_t^{s,x} = \sum_{j=0}^{2^{k-1}} I_{2^j}$, $k\geq 2$. We will carry out the computations for $I_{2^{k-1}}$, as it can be seen, all terms are treated analogously by choosing $j=1,\dots,2^{k-1}$. Then $I_{2^{k-1}}$ will take the following form

$$I^n_{2^{k-1}} = \sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}}\int_{\Delta^m}\mathcal{A}_k^X(u)du_1^k\cdots du^k_{m_k}\cdots du^1_1\cdots du^1_{m_1}$$

for the integrand

$$\begin{aligned}\mathcal{A}_k^X(u) := g(u_1^1)\cdots g(u^1_{r_1})\Bigg[g(u_1^2)\cdot g(u^2_{r_2})\Bigg[\cdots g(u_1^{k+1})\cdots\\ \cdots g(u^k_{m_k})\Bigg]g(u^{k-1}_{r_{k-1}+1})\cdots g(u^2_{m_2})\Bigg]g(u^1_{r_1+1})\cdots g(u^1_{m_1}),\end{aligned}$$

where the functions g denote elements in the set

$$g(u)\in\{Db(X_u^{s,x}), D^2b(X_u^{s,x}),\cdots, D^kb(X_u^{s,x})\}.$$

Let $p\in[1,\infty)$ choose $r,s\in[1,\infty)$ such that $sp=2^q$ for some integer q and $\frac{1}{r}+\frac{1}{s}=1$.

Then using Girsanov's theorem and Hölder's inequality we have

$$E\left[\|I_{2^{k-1}}\|^p\right] \leq CE\left[\left\| \sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}} \int_{\Delta^m} \mathcal{A}_k^{B^H}(u)du_1^k\cdots du_{m_k}^k\cdots du_1^1\cdots du_{m_1}^1\right\|^{2^q}\right]^{p/2^q}.$$

Now using the maximum norm on $\mathbb{R}^{kd\times d}$ we get

$$E\left[\|I_{2^{k-1}}\|^p\right] \\ \leq C\left(\sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}} \sum_{i,j,l_1,\dots,l_{k-1}=1}^{d} \left\| \int_{\Delta^m} \mathcal{B}_k^{B^H}(u)du_1^k\cdots du_{m_k}^k\cdots du_1^1\cdots du_{m_1}^1\right\|_{L^{2^q}(\Omega,\mathbb{R})}\right)^p, \tag{7.44}$$

where

$$\mathcal{B}_k^{B^H}(u) := h(u_1^1)\cdots h(u_{r_1}^1)\Big[h(u_1^2)\cdot h(u_{r_2}^2)\Big[\cdots h(u_1^{k+1})\cdots \\ \cdots h(u_{m_k}^k)\Big]h(u_{r_{k-1}+1}^{k-1})\cdots h(u_{m_2}^2)\Big]h(u_{r_1+1}^1)\cdots h(u_{m_1}^1),$$

where the functions h denote elements in the set

$$h(u) \in \left\{\frac{\partial^k}{\partial x_j\partial x_{l_1}\cdots\partial x_{l_{k-1}}}b^{(i)}(B_u^H),\right\}_{i,l_1,\dots,l_{k-1},j=1,\dots,d}.$$

As we saw at the beginning of the proof, the integral in (7.44) can be written as a sum like in (7.42) of integrals on simplices, that is

$$\int_{\Delta^m} \mathcal{B}_k^{B^H}(u)du = \sum_{\sigma^{-1}} \int_{\Delta_{s,t}^{m_1^+}} \prod_{l=1}^{m_1^+} f_{\sigma(l)}(w_l)dw, \quad f_l\in\Lambda, \quad l=1,\dots,m_1^+, \tag{7.45}$$

where

$$\Lambda := \left\{\frac{\partial^k b^{(i)}(B_\cdot^H)}{\partial x_j\partial x_{l_1}\cdots\partial x_{l_{k-1}}}; i,l_1,\dots,l_{k-1},j=1,\dots,d, r=1,\dots,m_1^+\right\}$$

and the number of terms in the sum in (7.45) is

$$\#\sum_{\sigma^{-1}} = \prod_{i=1}^{k}\frac{(r_i+m_i^+-1)!}{(r_i-1)!m_i^+!},$$

where $m_{k+1}^+ := 0$. It can be checked that for a sufficiently large enough constant $C>0$ independent of m we have

$$\prod_{i=1}^{k}\frac{(r_i+m_i^+-1)!}{(r_i-1)!m_i^+!} \leq C^{m_1^+}.$$

As a consequence

$$E\left[\|I_{2^{k-1}}\|^p\right] \\ \leq C\Bigg(\sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}} \sum_{i,j,l_1,\dots,l_{k-1}=1}^{d} \sum_{\sigma} \Big\| \int_{\Delta_{s,t}^{m_1^+}} \prod_{l=1}^{m_1^+} f_{\sigma(l)}(w_l)dw \Big\|_{L^{2^q}(\Omega,\mathbb{R})}\Bigg)^p, \quad f_l\in\Lambda. \tag{7.46}$$

Define

$$J := \int_{\Delta_{s,t}^{m_1^+}} \prod_{l=1}^{m_1^+} f_{\sigma(l)}(w)dw, \quad f_l\in\Lambda, \quad l=1,\dots,m_1^+. \tag{7.47}$$

Then using the same argument as in (7.34) by exploiting the identity in (7.5) repeatedly, we find that J can be written to the power 2 as a sum of, at most $2^{2m_1^+}$ of length $2m_1^+$ of the form

$$\int_{\Delta_{s,t}^{2m_1^+}} \prod_{l=1}^{2m_1^+} f_{\sigma(l)}(w)dw, \quad f_l\in\Lambda, \quad l=1,\dots,2m_1^+.$$

Repeating this argument, we find that we can write J^{2^q} as a sum of at most $2^{q2^q m_1^+}$ of length $2^q m_1^+$ of the form

$$\int_{\Delta_{s,t}^{2^q m_1^+}} \prod_{l=1}^{2^q m_1^+} f_{\sigma(l)}(w)dw, \quad f_l\in\Lambda, \quad l=1,\dots,2^q m_1^+.$$

Finally, taking expectation and choosing H small enough we can apply the estimate from Proposition 7.6 with $\kappa\equiv 1$ (or $\varepsilon_j=0$ for all j). Then we can find a constant $C_T>0$ such that

$$\begin{aligned}
E\|I_{2^{k-1}}\|^p &\leq \sum_{\substack{m\geq 1\\ r_l\leq m_l^-\\ l=1,\dots,k-1}} k^{m_1^+}(d)^{m_1^+ +1}\left(C_T^{2^q m_1^+}\frac{\|b\|_{L^1(\mathbb{R}^d)}^{2^q m_1^+}}{\Gamma\left((2^q m_1^+ +1)(1-d(2k+1)H)+1\right)^{1/2}}\right)^{1/2^q}\\
&\leq \sum_{m_1,\dots,m_k\geq 1} (m_1^+)^k k^{m_1^+}(d)^{m_1^+ +1} C_T^{m_1^+}\frac{\|b\|_{L^1(\mathbb{R}^d)}^{m_1^+}}{\Gamma\left((2^q m_1^+ +1)(1-d(2k+1)H)+1\right)^{1/2^{q+1}}}\\
&\leq \sum_{n\geq 1} n^k k^n (d)^{n+1} C_T^n \frac{\|b\|_{L^1(\mathbb{R}^d)}^n}{\Gamma\left((2^q n+1)(1-d(2k+1)H)+1\right)^{1/2^{q+1}}}\\
&<\infty
\end{aligned}$$

provided $H<\frac{1}{d(2k+1)}$.

As a result,

$$\sup_{s,t\in[0,T]}\sup_{x\in\mathbb{R}^d} E\left[\left\|\frac{\partial^k}{\partial x^k}X_t^{s,x}\right\|^p\right] \leq C(\|b\|_{L^1(\mathbb{R}^d)})$$

for a continuous function $C : [0,\infty) \to [0,\infty)$, depending on k, d, H, p and T. $\square$

Corollary 7.17. *Let $b \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Then for every $p \geq 1$ we have*

$$\sup_{s,t\in[0,T]} \sup_{x\in\mathbb{R}^d} E\left[\left\|\frac{\partial^k}{\partial x^k} X_t^{s,x}\right\|^p\right] \leq C(\|b\|_{L^1(\mathbb{R}^d)})$$

for a continuous function $C : [0,\infty) \to [0,\infty)$, depending on k, d, H, p and T provided $H < \min\left\{\frac{1}{2(3d-1)}, \frac{1}{d(2k+1)}\right\}$.

Proof. We need $H < \frac{1}{2(3d-1)}$ to ensure existence by Theorem 7.8. Then take $\{b_n\}_{n\geq1} \subset C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$ with $b_n \to b$ in $L^1(\mathbb{R}^d)$ as $n \to \infty$. The estimates obtained in Proposition 7.16 are independent of the size of the derivatives of b and hence the result follows. $\square$

The following is the main result of this section and shows that the fractional Brownian motion B^H creates a regularising effect on the solution as a function of the initial condition.

Theorem 7.18. *Assume $b \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Let $U \subset \mathbb{R}^d$ and open and bounded subset and $X = \{X_t, t \in [0,T]\}$ the solution of (7.1). Then for a small enough Hurst parameter $H < 1/2$ it follows*

$$X_t^{\cdot} \in \bigcap_{p>1} L^2(\Omega, W^{k,p}(U)).$$

Proof. First of all, approximate the irregular drift vector field b by a sequence of functions $b_n : \mathbb{R}^d \to \mathbb{R}^d$, $n \geq 1$ in $C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$ such that $b_n \to b$ in $L^1(\mathbb{R}^d)$ as $n \to \infty$. Denote by $X^{n,x} = \{X_t^{n,x}, t \in [0,T]\}$, the corresponding solution to (7.1) starting from $x \in \mathbb{R}^d$ when b is replaced by b_n.

Observe that for any test function $\varphi \in C_0^\infty(U, \mathbb{R}^d)$ and fixed $t \in [0,T]$ the set of random variables

$$\langle X_t^{n,\cdot}, \varphi\rangle := \int_U \langle X_t^{n,x}, \varphi(x)\rangle_{\mathbb{R}^d} dx, \quad n \geq 1$$

is relatively compact in $L^2(\Omega)$. To show this, we use the compactness criterion from Appendix, in Corollary 7.21 in terms of the Malliavin derivative. Since the Malliavin derivative is a closed linear operator we have

$$\begin{aligned} E[\int_0^T |D_\theta^j \langle X_t^{n,\cdot}, \varphi\rangle|^2 ds] &= \sum_{i=1}^d \left(\int_U E[D_\theta^j X_t^{n,x,(i)}]\varphi_i(x) dx\right)^2 \\ &\leq d\|\varphi\|^2_{L^2(\mathbb{R}^d,\mathbb{R}^d)} \lambda\{\text{supp }(\varphi)\} \sup_{x\in U} E\left[\int_0^T \|D_\theta X_t^{n,x}\|^2 ds\right], \end{aligned}$$

where D^j denotes the Malliavin derivative in the direction of $B^{H,(j)}$, λ the Lebesgue measure on $\mathbb{R}^d$, supp (φ) the support of φ and $\|\cdot\|$ a matrix norm. Then taking sum over all $j = 1, \ldots, d$ and using Lemma 7.14 we obtain

$$\sup_{n\geq1} \|D_\cdot \langle X_t^{n,\cdot}, \varphi\rangle\|^2_{L^2(\Omega\times[0,T])} \leq C\|\varphi\|^2_{L^2(\mathbb{R}^d,\mathbb{R}^d)} \lambda\{\text{supp }(\varphi)\}.$$

In a similar manner we have

$$\sup_{n\geq 1}\int_0^T\int_0^T \frac{E[\|D_{\theta'}\langle X_t^{n,\cdot},\varphi\rangle - D_\theta\langle X_t^{n,\cdot},\varphi\rangle\|^2]}{|\theta'-\theta|^{1+2\beta}} < \infty$$

for $\beta\in(0,1/2)$. Hence $\langle X_t^{n,\cdot},\varphi\rangle$, $n\geq 1$ is relatively compact in $L^2(\Omega)$. Let us denote by $Y_t(\varphi)$ its limit after taking (if necessary) a subsequence.

Following exactly the same reasoning as in Lemma 7.13 one can show that

$$\langle X_t^{n,\cdot},\varphi\rangle \xrightarrow{n\to\infty} \langle X_t^{\cdot},\varphi\rangle$$

weakly in $L^2(\Omega)$. Then by uniqueness of the limit we can establish that

$$Y_t(\varphi) \overset{L^2(\Omega)}{=} \langle X_t^{\cdot},\varphi\rangle.$$

Note that there exists a subsequence $n(j)$ such that $\langle X_t^{n(j),\cdot},\varphi\rangle$ converges for every φ, that is, $n(j)$ is independent of φ.

We have that $X_t^{n,\cdot}$ is bounded in the Sobolev norm $L^2(\Omega, W^{k,p}(U))$ for each $n\geq 1$. Indeed, by Proposition 7.16 we have for a small enough $H<1/2$

$$\begin{aligned}\sup_{n\geq 1}\|X_t^{n,\cdot}\|^2_{L^2(\Omega,W^{k,p}(U))} =&\sup_{n\geq 1}\sum_{i=0}^k E\left[\|\frac{\partial^i}{\partial x^i}X_t^{n,\cdot}\|^2_{L^p(U)}\right]\\ \leq&\sum_{i=0}^k\int_U \sup_{n\geq 0} E\left[\|\frac{\partial^i}{\partial x^i}X_t^{n,x}\|^p\right]dx\\ <&\infty.\end{aligned}$$

Since $L^2(\Omega, W^{k,p}(U))$, $p\in(1,\infty)$ is reflexive, by Banach-Alaoglu's theorem we get that the set $\{X_t^{n,x}\}_{n\geq 1}$ is weakly compact in the $L^2(\Omega, W^{k,p}(U))$-topology. Thus, there exists a subsequence $n(j)$, $j\geq 0$ such that

$$X_t^{n(j),\cdot} \xrightarrow[j\to\infty]{w} Y \in L^2(\Omega, W^{k+1,p}(U)).$$

On the other hand, we have proven that $X_t^{n,x}\to X_t^x$ strongly in $L^2(\Omega)$, so by uniqueness of the limit we can conclude that

$$X_t^{\cdot} = Y \in L^2(\Omega, W^{k,p}(U)),\quad P-a.s.$$

Moreover, for all $A\in\mathcal{F}$ and $\varphi\in C_0^\infty(\mathbb{R}^d,\mathbb{R}^d)$ we have

$$E[1_A\langle X_t^{\cdot},\varphi'\rangle] = \lim_{j\to\infty} E[1_A\langle X_t^{n(j),\cdot},\varphi'\rangle] = \lim_{j\to\infty} -E[1_A\langle \frac{\partial}{\partial x}X_t^{n(j),\cdot},\varphi\rangle] = -E[1_A\langle Y',\varphi\rangle]$$

and thus

$$\langle X_t^{\cdot},\varphi'\rangle = -\langle Y',\varphi\rangle,\quad P-a.s.$$

□

Appendix

7.A Technical results

The following result which is due to [30, Theorem 1] provides a compactness criterion for subsets of $L^2(\Omega)$ using Malliavin calculus.

Theorem 7.19. *Let $\{(\Omega, \mathcal{A}, P)\,; H\}$ be a Gaussian probability space, that is $(\Omega, \mathcal{A}, P)$ is a probability space and H a separable closed subspace of Gaussian random variables of $L^2(\Omega)$, which generate the σ-field $\mathcal{A}$. Denote by $\mathbf{D}$ the derivative operator acting on elementary smooth random variables in the sense that*

$$\mathbf{D}(f(h_1, \ldots, h_n)) = \sum_{i=1}^{n} \partial_i f(h_1, \ldots, h_n) h_i,\ h_i \in H, f \in C_b^\infty(\mathbb{R}^n).$$

Further let $\mathbb{D}^{1,2}$ be the closure of the family of elementary smooth random variables with respect to the norm

$$\|F\|_{1,2} := \|F\|_{L^2(\Omega)} + \|\mathbf{D}F\|_{L^2(\Omega;H)}\,.$$

Assume that C is a self-adjoint compact operator on H with dense image. Then for any $c > 0$ the set

$$\mathcal{G} = \left\{G \in \mathbb{D}^{1,2} : \|G\|_{L^2(\Omega)} + \left\|C^{-1}\mathbf{D}\,G\right\|_{L^2(\Omega;H)} \leq c\right\}$$

is relatively compact in $L^2(\Omega)$.

In order to formulate compactness criteria useful for our purposes, we need the following technical result which also can be found in [30].

Lemma 7.20. *Let $v_s, s \geq 0$ be the Haar basis of $L^2([0,T])$. For any $0 < \alpha < 1/2$ define the operator A_α on $L^2([0,T])$ by*

$$A_\alpha v_s = 2^{k\alpha} v_s, \text{ if } s = 2^k + j$$

for $k \geq 0, 0 \leq j \leq 2^k$ and

$$A_\alpha 1 = 1.$$

Then for all β with $\alpha < \beta < (1/2)$, there exists a constant c_1 such that

$$\|A_\alpha f\| \leq c_1 \left\{ \|f\|_{L^2([0,T])} + \left(\int_0^T \int_0^T \frac{|f(t) - f(t')|^2}{|t - t'|^{1+2\beta}} dt\, dt' \right)^{1/2} \right\}.$$

A direct consequence of Theorem 7.19 and Lemma 7.20 is now the following compactness criteria.

Corollary 7.21. *Let a sequence of $\mathcal{F}_T$-measurable random variables $X_n \in \mathbb{D}^{1,2}$, $n = 1, 2...$, be such that there exists a constant $C > 0$ with*

$$\sup_n E[|X_n|^2] \le C,$$

$$\sup_n E\left[\|D_t X_n\|^2_{L^2([0,T])}\right] \le C$$

and there exists a $\beta \in (0, 1/2)$ such that

$$\sup_n \int_0^T \int_0^T \frac{E\left[\|D_t X_n - D_{t'} X_n\|^2\right]}{|t - t'|^{1+2\beta}} dt dt' < \infty$$

where $\|\cdot\|$ denotes any matrix norm.

Then the sequence X_n, $n = 1, 2...$, is relatively compact in $L^2(\Omega)$.

For the use of the above result we will need to exploit the following technical results.

Lemma 7.22. *Let $H \in (0, 1/2)$ and $s, t \in [0, T]$, $s < t$. Then*

$$\int_0^t \int_0^t \frac{|K_H(t, \theta') - K_H(t, \theta)|^2}{|\theta' - \theta|^\gamma} d\theta d\theta' < \infty \tag{7.48}$$

for $\gamma \in (1, 2)$.

Proof. Write

$$K_H(t, \theta') - K_H(t, \theta) = c_H \left[f_t(\theta') - f_t(\theta) + \left(H - \frac{1}{2}\right)(g_t(\theta) - g_t(\theta')) \right],$$

where $f_t(s) := \left(\frac{t}{s}\right)^{H-\frac{1}{2}} (t - s)^{H-\frac{1}{2}}$ and $g_t(s) := s^{\frac{1}{2}-H} \int_s^t u^{H-\frac{3}{2}} (u - s)^{H-\frac{1}{2}} du$. Hence,

$$\begin{aligned}
\int_0^t \int_0^t \frac{|K_H(t, \theta') - K_H(t, \theta)|^2}{|\theta' - \theta|^\gamma} d\theta d\theta' = &\int_0^t \int_0^{\theta'} \frac{(K_H(t, \theta') - K_H(t, \theta))^2}{(\theta' - \theta)^\gamma} d\theta d\theta' \\
&+ \int_0^t \int_{\theta'}^t \frac{(K_H(t, \theta) - K_H(t, \theta'))^2}{(\theta - \theta')^\gamma} d\theta d\theta' \\
=:& I_1 + I_2,
\end{aligned}$$

where I_1 and I_2 the respective integrals. One can observe that the challenge is to compute either integral. So we will just show the computations for I_1.

We have

$$\begin{aligned}
I_1 \le& c_H^2 \left[\int_0^t \int_0^{\theta'} \frac{(f_t(\theta) - f_t(\theta'))^2}{(\theta' - \theta)^\gamma} d\theta d\theta' + \left(H - \frac{1}{2}\right)^2 \int_0^t \int_0^{\theta'} \frac{(g_t(\theta) - g_t(\theta'))^2}{(\theta' - \theta)^\gamma} d\theta d\theta' \right] \\
=:& c_H^2 \left(I_1^1 + \left(H - \frac{1}{2}\right)^2 I_1^2 \right),
\end{aligned}$$

where I_1^1 and I_1^2 are the respective integrals above.

Now for I_1^2 one can show that $g_t(\theta)$ is Hölder-continuous of degree $1/2$ on $[0,1]$, hence

$$|g_t(\theta) - g_t(\theta')| \leq C|\theta - \theta'|^{1/2}$$

for all $t \in [0,T]$ and therefore

$$\int_0^t \int_0^{\theta'} \frac{(g_t(\theta) - g_t(\theta'))^2}{(\theta' - \theta)^\gamma} d\theta d\theta' < \infty$$

for $\gamma \in (1,2)$.

For I_1^2 we have

$$I_1^1 \leq \sup_{\theta' \in (0,t)} \left(\frac{t}{\theta'}\right)^{2H-1} \int_0^t \int_0^{\theta'} \frac{((t-\theta)^{H-\frac{1}{2}} - (t-\theta')^{H-\frac{1}{2}})^2}{(\theta'-\theta)^\gamma} d\theta d\theta'.$$

Now apply the change of variables $v = t - \theta$ and $wv = \theta' - \theta$ in order to get

$$\begin{aligned} &\int_0^t \int_0^{\theta'} \frac{((t-\theta)^{H-\frac{1}{2}} - (t-\theta')^{H-\frac{1}{2}})^2}{(\theta'-\theta)^\gamma} d\theta d\theta' \\ &\quad = \int_0^t v^{2\left(H-\frac{1}{2}\right)-\gamma+1} \int_0^1 \frac{(1-(1-w)^{H-\frac{1}{2}})^2}{w^\gamma} dw dv. \end{aligned}$$

By using standard techniques one can prove that

$$C := \int_0^1 \frac{(1-(1-w)^{H-\frac{1}{2}})^2}{w^\gamma} dw < \infty.$$

for any $\gamma < 2$. Finally, integrating with respect to v we obtain

$$\int_0^t \int_0^{\theta'} \frac{((t-\theta)^{H-\frac{1}{2}} - (t-\theta')^{H-\frac{1}{2}})^2}{(\theta'-\theta)^\gamma} d\theta d\theta' = C \int_0^t v^{2\left(H-\frac{1}{2}\right)-\gamma+1} dv < \infty$$

provided $\gamma < 2H + 1$, i.e. $\gamma \in (1,2)$. □

Lemma 7.23. *Let $H \in (0,1/2)$, $w_j > -1$, $j = 1, \dots, 2m$, $\theta, t \in [0,T]$, $\theta < t$ and $(\varepsilon_1, \dots, \varepsilon_{2m}) \in \{0,1\}^{2m}$ be fixed. Then exists a finite constant $C > 0$ such that*

$$\begin{aligned} &\int_{\Delta_{\theta,t}^{2m}} \prod_{j=1}^{2m} (K_H(s_j,\theta) - K_H(s_j,\theta'))^{\varepsilon_j} |s_j - s_{j-1}|^{w_j} ds \\ &\quad \leq C \frac{\prod_{j=1}^{2m} \Gamma(w_j+1) \, |\theta'-\theta|^{\gamma \sum_{j=1}^{2m} \varepsilon_j} |t-\theta|^{2m + \sum_{j=1}^{2m} w_j + (H-\frac{1}{2}-\gamma)\sum_{j=1}^{2m} \varepsilon_j}}{\Gamma\left(2m+1+\sum_{j=1}^{2m} w_j + (H-\frac{1}{2}-\gamma)\sum_{j=1}^{2m} \varepsilon_j\right)} \end{aligned}$$

for $\gamma \in (0,H)$. Observe that if $\varepsilon_j = 0$ for all $j = 1, \dots, 2m$ we obtain the classical formula.

Proof. First, observe that given exponents $a, b > -1$ we have

$$\int_{\theta}^{s_{j+1}} (s_{j+1}-s_j)^a (s_j-\theta)^b ds_j = \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+b+2)} (s_{j+1}-\theta)^{a+b+1}.$$

Let us start computing

$$\int_{\theta}^{s_2} (K_H(s_1,\theta)-K_H(s_1,\theta'))^{\varepsilon_1} |s_2-s_1|^{w_2} |s_1-\theta|^{w_1} ds_1$$

for $w_1, w_2 \geq -1$.

Define the functions $f_t(s)$ and $g_t(s)$, $s,t \in [0,T]$, $s \leq t$ as in Lemma 7.22. Then for some finite constant $C_H > 1$

$$(K_H(s_1,\theta)-K_H(s_1,\theta'))^{\varepsilon_1} \leq C_H \Bigg((f_{s_1}(\theta)-f_{s_1}(\theta'))^{\varepsilon_1} + (g_{s_1}(\theta')-g_{s_1}(\theta))^{\varepsilon_1} \Bigg).$$

For the second term we simply have

$$|g_{s_1}(\theta')-g_{s_1}(\theta)|^{\varepsilon_1} \leq C|\theta'-\theta|^{\varepsilon_1/2}.$$

since g_{s_1} is Hölder continuous of order $1/2$ uniformly in s_1. So

$$\begin{aligned}\int_{\theta}^{s_2} &(g_{s_1}(\theta)-g_{s_1}(\theta'))^{\varepsilon_1} |s_2-s_1|^{w_2} |s_1-\theta|^{w_1} ds_1 \\ &\leq C|\theta-\theta'|^{\varepsilon_1/2} \frac{\Gamma(w_1+1)\Gamma(w_2+1)}{\Gamma(w_1+w_2+2)} |s_2-\theta|^{w_1+w_2+1}.\end{aligned}$$

For the term depending on f_s, as before, observe that

$$\frac{y^{-\alpha}-x^{-\alpha}}{(x-y)^{\gamma}} \leq C y^{-\alpha-\gamma}$$

for every $0 < y < x < T$ and $\alpha := (\frac{1}{2}-H) \in (0,1/2)$ and $\gamma < \frac{1}{2}-\alpha$. Hence

$$\begin{aligned}\int_{\theta}^{s_2} &(f_{s_1}(\theta)-f_{s_1}(\theta'))^{\varepsilon_1} |s_2-s_1|^{w_2} |s_1-\theta|^{w_1} ds_1 \\ \leq& C|\theta-\theta'|^{\gamma\varepsilon_1} \int_{\theta}^{s_2} (s_2-s_1)^{w_2} (s_1-\theta)^{w_1-(\alpha+\gamma)\varepsilon_1} ds_1 \\ =& C|\theta-\theta'|^{\gamma\varepsilon_1} \frac{\Gamma(w_1-(\alpha+\gamma)\varepsilon_1+1)\Gamma(w_2+1)}{\Gamma(w_1+w_2-(\alpha+\gamma)\varepsilon_1+2)} |s_2-\theta|^{w_1+w_2-(\alpha+\gamma)\varepsilon_1+1}.\end{aligned}$$

Observe that

$$\frac{\Gamma(w_1-(\alpha+\gamma)\varepsilon_1+1)\Gamma(w_2+1)}{\Gamma(w_1+w_2-(\alpha+\gamma)\varepsilon_1+2)} \geq \frac{\Gamma(w_1+1)\Gamma(w_2+1)}{\Gamma(w_1+w_2+2)}$$

for all $\alpha \in (0, 1/2)$, $\gamma \in (0, \frac{1}{2} - \alpha)$, $w_1, w_2 > -1$ and $\varepsilon_1 \in \{0, 1\}$. Altogether we get

$$\int_\theta^{s_2} (K_H(s_1, \theta) - K_H(s_1, \theta'))^{\varepsilon_1} |s_2 - s_1|^{w_2} |s_1 - \theta|^{w_1} ds_1$$
$$\leq C_{H,T} |\theta - \theta'|^{\gamma\varepsilon_1} \frac{\Gamma(w_1 - (\alpha+\gamma)\varepsilon_1 + 1)\,\Gamma(w_2+1)}{\Gamma(w_1 + w_2 - (\alpha+\gamma)\varepsilon_1 + 2)} |s_2 - \theta|^{w_1 + w_2 - (\alpha+\gamma)\varepsilon_1 + 1}.$$

Integrating iteratively we obtain the desired formula. □

Finally, we give a similar estimate which is used in Lemma 7.14.

Lemma 7.24. *Let $H \in (0, 1/2)$, $w_j > -1$, $j = 1, \dots, 2m$, $\theta, t \in [0, T]$, $\theta < t$ and $(\varepsilon_1, \dots, \varepsilon_{2m}) \in \{0, 1\}^{2m}$ be fixed. Then exists a finite constant $C > 0$ such that*

$$\int_{\Delta^{2m}_{\theta,t}} \prod_{j=1}^{2m} (K_H(s_j, \theta))^{\varepsilon_j} |s_j - s_{j-1}|^{w_j} ds$$
$$\leq C \frac{\prod_{j=1}^{2m} \Gamma(w_j + 1)\, |t - \theta|^{2m + \sum_{j=1}^{2m} w_j + (H - \frac{1}{2}) \sum_{j=1}^{2m} \varepsilon_j}}{\Gamma\left(2m + 1 + \sum_{j=1}^{2m} w_j + (H - \frac{1}{2}) \sum_{j=1}^{2m} \varepsilon_j\right)}.$$

Observe that if $\varepsilon_j = 0$ for all $j = 1, \dots, 2m$ we obtain the classical formula.

Proof. Let us start computing

$$\int_\theta^{s_2} (K_H(s_1, \theta))^{\varepsilon_1} |s_2 - s_1|^{w_2} |s_1 - \theta|^{w_1} ds_1$$

for $w_1, w_2 \geq -1$.

Define the functions $f_t(s)$ and $g_t(s)$, $s, t \in [0, T]$, $s \leq t$ as in the proof of Lemma 7.22. Then for some finite constant $C_{H,T} > 0$

$$|K_H(s_1, \theta)|^{\varepsilon_1} ds \leq C_{H,T} \left(|f_{s_1}(\theta)|^{\varepsilon_1} + |g_{s_1}(\theta)|^{\varepsilon_1}\right) \leq C_{H,T} \left(|f_{s_1}(\theta)|^{\varepsilon_1} + |\theta|^{\varepsilon_1/2}\right).$$

Then we have

$$\int_\theta^{s_2} (K_H(s_1, \theta))^{\varepsilon_1} |s_2 - s_1|^{w_2} |s_1 - \theta|^{w_1} ds_1$$
$$\leq C_{H,T} \frac{\Gamma\left(w_1 + (H - \frac{1}{2})\varepsilon_1 + 1\right) \Gamma(w_2 + 1)}{\Gamma\left(w_1 + w_2 + (H - \frac{1}{2})\varepsilon_1 + 2\right)} |s_2 - \theta|^{w_1 + (H - \frac{1}{2})\varepsilon_1 + 1}.$$

Integrating iteratively one obtains the desired estimate. □

Chapter 8

Future work

The overall aim of this thesis has been to investigate in detail the characteristics of stochastic differential equations driven by Brownian motion and fractional Brownian motion. We have answered some of the questions that were posed. Also, we have found an application in mathematical finance where our findings can be applied notwithstanding, a good deal of new directions and problems can be considered from now on.

Let us briefly give some very short ideas for future development and understanding of stochastic differential equations. Let B^H be a d-dimensional fractional Brownian motion with Hurst parameter $H \in (0,1)$, defined on a probability space $(\Omega, \mathcal{F}, P)$ equipped with the filtration $\{\mathcal{F}_t\}_{T\in[0,T]}$ generated by B^H and augmented by all P-null sets. Let X be the process defined as the unique (global) strong solution of

$$X_t = x + \int_0^t b(s, X_s)ds + B_t^H, \quad t \in [0,T] \tag{8.1}$$

where x is an initial condition in $\mathbb{R}^d$ and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ a vector field for which one has strong existence and uniqueness.

8.1 Construction of solutions of stochastic differential equations and invariant manifolds

We have seen the strength of the method employed to construct strong solutions of stochastic differential equations. As it was the case in Chapter 7 when the driving noise was neither Markovian nor a semimartingale. Nevertheless, these are not the only cases one can study. There is a considerable amount of different types of equations one can look at. For example, equations of the form

$$dX_t = dA_t + dB_t^H,$$

where A is a process of bounded variation which arises from limits of the form

$$\lim_{n\to\infty} \int_0^t b_n(X_s)ds$$

for coefficients b_n, $n \geq 1$. Some findings in this direction have been attained in [16] for the classical case of Hurst parameter $1/2$.

Also, in infinite dimensions, i.e.

$$dX_t = (AX_t + b(X_t))dt + QdW_t^H$$

for (mild) solutions X, where A i a densely defined operator A of parabolic type on a separable Hilbert space $\widetilde{H}$, $b : \widetilde{H} \to \widetilde{H}$ is an irregular functional, Q a Hilbert-Schmidt operator and $W_\cdot$ a cylindrical fractional Brownian motion. In the classical Brownian motion case the authors in [48] proved existence and Malliavin differentiability of the solution when b is Hölder continuous.

Another prominent line of investigation that could be considered is to establish the existence of a finite Lyapunov spectrum and invariant (Sobolev) manifolds for d-dimensional stochastic differential equations of the form

$$dX_t = b(X_t)dt + dB_t^H, \quad t \geq 0, \quad X_0 = x \in \mathbb{R}^d \tag{8.2}$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a Borel measurable and locally Lebesgue integrable vector field (with no presumed differentiability or even continuity requirements). The driving noise B^H is d-dimensional fractional Brownian motion on a complete Wiener space $(\Omega, \mathcal{F}, P)$, where P is the Wiener measure on the space Ω of all continuous paths $\mathbb{R} \to \mathbb{R}^d$ given the compact open topology and the P-complete Borel σ-algebra $\mathcal{F}$. See e.g. [68], where such singular models are e.g. discussed in connection with problems in statistical mechanics.

A recent work in [87] is focused on the existence of a Sobolev differentiable stochastic flow for the singular equation (8.2) when the driving noise is taken to be a standard Brownian motion, i.e., $H = \frac{1}{2}$. A substantial difference is that their method relies on the fact that when $H = 1/2$ the driving noise is a Markovian martingale contrary to the case $H \neq 1/2$ where B_t^H is neither a semimartingale nor a Markovian process which brings additional challenges to the problem, as for instance no PDE approach can be used. The approach adopted in this work provides a unique and somewhat surprising perspective to the existing theory of finite-dimensional stochastic (and deterministic) dynamical systems: well-posedness of the initial value problem (8.2) driven by bounded and/ or integrable measurable drift vector field b. No regularity or even continuity hypotheses are imposed on the driving vector field b. Furthermore, under these hypotheses it is possible to construct a unique stochastic flow of Sobolev diffeomorphisms for the equation (8.2). Then the question whether one can extend it to any arbitrary Hurst parameter H remains open and it appears to be highly non-trivial. In this thesis we answered the question on existence and uniqueness when b is bounded and integrable.

8.2 Regularity of densities

There is still an enormous amount of work to be investigated in the topic of densities of solutions with irregular coefficients. For instance, in equation (8.1) in the case $H = 1/2$ it was shown in Chapter 2 that if $b \in W^{1,\infty}(\mathbb{R}^d)$ then $X_t \in \mathbb{D}^{2,p}$ for any $p \geq 1$. This was shown to be optimal in the sence that $X_t \notin \mathbb{D}^{3,p}$ for any $p \neq 1$. We know in view of the results in [5] that nondeterministic random variables in $\cap_{p\geq 1}\mathbb{D}^{2,p}$ with invertible Malliavin covariance matrix admit a Hölder continuous density with order $\alpha \in (0,1)$. Nevertheless, this does not seem to be an optimal result because if we look at the equation

$$dX_t = \text{sign}\,(X_t)dt + dB_t, \quad X_0 = x \in \mathbb{R}^d, \quad t \geq 0$$

where sign $(x) = \frac{x}{|x|}1_{\{x\neq 0\}}$, $x \in \mathbb{R}^d$ denotes the generalised *signum* function, then it was shown in Chapter 4, at least in dimension one, that the density of X_t at any time $t > 0$ is globally Lipschitz continuous in spite of the drift being discontinuous and in spite of X_t not being twice Malliavin differentiable for any $t > 0$. Does this give a hint that once Malliavin differentiability and probably some technical conditions imply that the density is continuous? This is indeed not the case in the space $\mathbb{D}^{1,1}$ where one can both find random variables with Lipschitz densities and without densities at all. Another example is the fractional space $\mathbb{D}^{s,2}$, $0 < s < 1/2$. For instance, it was proven in [92] that the local-time of a standard Brownian motion, denote it by L_t, belongs to the space $L_t \in \mathbb{D}^{1/2-,p}$ for any $p \geq 1$ and $L_t \notin \mathbb{D}^{1/2,p}$ for any $p \geq 1$ while its density is C^∞. For this reason, it remains unclear what properties of the densities of random variables belonging to these spaces one may expect.

Let us consider an Itô process as in Chapter 4 of the form

$$X_u(t) = x + \int_0^t u(s)ds + B(t), \quad t \geq 0$$

where B is a standard d-dimensional Brownian motion and u is a bounded, adapted process with integrable trajectories. We have very recently proven in [13] that the densities of $X(t)$ at any given time $t > 0$ are Hölder continuous of any order even if u is merely bounded when the dimension is one. By exploiting the same method one may be able to control the characteristic function. Indeed, a general approach to study the regularity of the densities is to find tail estimates of the Fourier-Stiltjes transform of the law of the random variable of interest. We showed sharp estimates of the characteristic function in dimension one by posing a control problem based on very similar ideas as in Chapter 4. The latter implies that the fundamental solution of the Fokker-Planck equation is in fact continuous which has been an open question. We also see that the high dimensional problem is more demanding and that the Fourier method is not conclusive in this case and hence the question whether the densities are continuous for $d \geq 2$ or whether there is a counterexample remains open.

8.3 Application to the sensitivity analysis

In Chapter 6 we have given generalisations of the results by [50] to discontinuous unbounded drift coefficients and pay-off functions and computed some *Deltas* of lookback and Asian options. We have given some approximation formulas for Asian option. However, the computation of sensitivities of Asian options appears to be a laborious task since well-known formulas involve Skorokhod-type expressions which are in general very difficult to simulate or simplify and require higher order of Malliavin smoothness. Therefore, it remains still open to find closed form expressions for the *Deltas* of Asian options where the underlying process is driven by some stochastic differential equation with very irregular coefficients. Another remaining improvement of our results is to generalise these formulas to higher dimensions although we are confident that the same approach can be done in higher dimensions at the expense of not having a closed explicit expression for the Malliavin derivative of the solution and the first variation process.

One can also look at non-Markovian models like for instance by introducing some delay in the price dynamics of the underlying stock dynamics. In other words, one can consider a *stochastic functional differential equation* of *delay*-type of the form

$$\begin{aligned} dx(t) &= f(t,x(t),x_t)dt + g(t,x(t),x_t)dB(t) + \int_{\mathbb{R}_0} h(t,x(t),x_t,z)\tilde{N}(dt,dz), \\ x_0 &= \eta \in L^2([-r,0];\mathbb{R}^d) \end{aligned}$$

where f,g and h are suitable functionals, B is a standard Brownian motion and $\tilde{N}$ a *compensated Poisson random measure*. Here, $r>0$ is some fixed time delay and x_t denotes the whole path of the process x from $t-r$ to t. The initial condition becomes then a whole deterministic function on the interval $[-r,0]$. For this reason, the study of the sensitivity becomes a different problem since now, the future values of the process x are not subject to a single initial value but a whole function. Then one is compelled to work in an infinite-dimensional setting and define the corresponding *Delta*-index as a functional derivative

$$\Delta := \frac{\partial}{\partial \eta} E[\Phi(x^\eta(t), x_t^\eta)]$$

where $\Phi : \mathbb{R}\times\mathbb{R} \to [0,\infty)$ is a pay-off function and here $\frac{\partial}{\partial\eta}$ denotes differentiation in the Fréchet sense. This problem has been already studied in [9] in the case of smooth coefficients f and g in the case of no jumps, i.e. $h=0$ and in [8] in the case of jumps.

Bibliography

[1] E. Alòs, O. Mazet, and D. Nualart, *Stochastic calculus with respect to Gaussian processes*, Ann. Probab. **29** (2001), no. 2, 766–801. MR 1849177 (2002g:60083)

[2] L. Ambrosio, *Transport equation and Cauchy problem for BV vector fields and applications*, Journées “Équations aux Dérivées Partielles”, École Polytech., Palaiseau, 2004, pp. Exp. No. I, 11. MR 2135356 (2006m:35030)

[3] D. Applebaum, *Lévy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009. MR 2512800 (2010m:60002)

[4] T. Atsushi, *Bismut-Elworthy-Li formulae for stochastic differential equations with jumps*, Journal of Theoretical Probability **23** (2004), no. 2, 576–604.

[5] V. Bally and L. Caramellino, *Riesz transform and integration by parts formulas for random variables*, Stochastic Process. Appl. **121** (2011), no. 6, 1332–1355. MR 2794979 (2012f:60182)

[6] V. Bally and A. Kohatsu-Higa, *Lower bounds for densities of Asian type stochastic differential equations*, J. Funct. Anal. **258** (2010), no. 9, 3134–3164. MR 2595738 (2010m:60190)

[7] D. Baños, *The Bismut-Elworthy-Li formula for mean-field stochastic differential equations*, (preprint 2015).

[8] D. Baños, F. Cordoni, G. Di Nunno, L. Di Persio, and E. E. Røse, *Stochastic systems with memory and jumps*, (work in progress).

[9] D. Baños, G. Di Nunno, and F. Proske, *Sensitivity analysis in a market with memory*, (preprint 2015).

[10] D. Baños, S. Duedahl, T. Meyer-Brandis, and F. Proske, *Computing Deltas without derivatives*, (preprint 2015).

[11] ———, *Construction of Malliavin differentiable strong solutions of SDEs under an integrability condition on the drift without the Yamada-Watanabe principle*, (preprint 2015).

[12] D. Baños and P. Krühner, *Optimal bounds for the densities of solutions of SDEs with measurable and path dependent drift coefficients*, arXiv:1408.2386 [math.PR].

[13] D. Baños and P. Krühner, *Hölder continuous densities of solutions of SDEs with measurable and path dependent drift coefficients. Regularity of the Fokker-Planck equation.*, (work in progress).

[14] D. Baños and T. Nilssen, *Malliavin and flow regularity of SDEs. Application to the study of densities and the stochastic transport equation*, arXiv:1410.0786 [math.PR].

[15] D. Baños, T. Nilssen, and F. Proske, *Strong existence and higher order Fréchet differentiability of stochastic flows of fractional Brownian motion driven SDE's with singular drift*, (preprint 2015).

[16] R. F. Bass and Z-Q. Chen, *Brownian motion with singular drift*, Ann. Probab. **31** (2003), no. 2, 791–817. MR 1964949 (2004b:60144)

[17] V. E. Beneš, *Girsanov functionals and optimal bang-bang laws for final value stochastic control*, Stochastic Processes Appl. **2** (1974), 127–140. MR 0363613 (50 #16051)

[18] E. Benhamou, *Optimal Malliavin weighting function for the computation of the Greeks*, Math. Finance **13** (2003), no. 1, 37–53, Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001). MR 1968095

[19] F. E. Benth, M. Groth, and O. Wallin, *Derivative-free Greeks for the Barndorff-Nielsen and Shephard stochastic volatility model*, Stochastics **82** (2010), no. 1-3, 291–313. MR 2677550 (2011k:60191)

[20] F. E. Benth, J. Šaltytė Benth, and S. Koekebakker, *Stochastic modelling of electricity and related markets*, Advanced Series on Statistical Science & Applied Probability, 11, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008. MR 2416686 (2009b:91001)

[21] J-M. Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston, Inc., Boston, MA, 1984. MR 755001 (86f:58150)

[22] A. N. Borodin and P. Salminen, *Handbook of Brownian motion—facts and formulae*, second ed., Probability and its Applications, Birkhäuser Verlag, Basel, 2002. MR 1912205 (2003g:60001)

[23] N. Bouleau and F. Hirsch, *Propriétés d'absolue continuité dans les espaces de Dirichlet et application aux équations différentielles stochastiques*, Séminaire de Probabilités, XX, 1984/85, Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 131–161. MR 942020 (89f:60060)

[24] ______, *Dirichlet forms and analysis on Wiener space*, de Gruyter Studies in Mathematics, vol. 14, Walter de Gruyter & Co., Berlin, 1991. MR 1133391 (93e:60107)

[25] R. Buckdahn, J. Li, and S. Peng, *Mean-field stochastic differential equations and associated pdes*, arXiv:1407.1215 [math.PR].

[26] P. Cheridito, *Arbitrage in fractional Brownian motion models*, Finance Stoch. **7** (2003), no. 4, 533–553. MR 2014249 (2004m:60076)

[27] R. Cont and D-A. Fournié, *Functional Itô calculus and stochastic integral representation of martingales*, Ann. Probab. **41** (2013), no. 1, 109–133. MR 3059194

[28] G. Da Prato, F. Flandoli, E. Priola, and M. Röckner, *Strong uniqueness for stochastic evolution equations in Hilbert spaces perturbed by a bounded measurable drift*, Ann. Probab. **41** (2013), no. 5, 3306–3344. MR 3127884

[29] G. Da Prato, F. Flandoli, M. Röckner, and A. Veretennikov, *Strong uniqueness for SDEs in Hilbert spaces with non-regular drift*, To appear in Annals of Probability, arXiv:1407.1215 [math.PR].

[30] G. Da Prato, P. Malliavin, and D. Nualart, *Compact families of Wiener functionals*, C. R. Acad. Sci. Paris Sér. I Math. **315** (1992), no. 12, 1287–1291. MR 1194537 (94a:60085)

[31] S. De Marco, *Smoothness and asymptotic estimates of densities for SDEs with locally smooth coefficients and applications to square root-type diffusions*, Ann. Appl. Probab. **21** (2011), no. 4, 1282–1321. MR 2857449 (2012g:60186)

[32] A. Debussche and N. Fournier, *Existence of densities for stable-like driven SDE's with Hölder continuous coefficients*, J. Funct. Anal. **264** (2013), no. 8, 1757–1778. MR 3022725

[33] L. Decreusefond and A. S. Üstünel, *Stochastic analysis of the fractional Brownian motion*, Potential Anal. **10** (1999), no. 2, 177–214. MR 1677455 (2000b:60133)

[34] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Ann. **300** (1994), no. 3, 463–520. MR 1304434 (95m:90022b)

[35] L. Delong, *Pricing and hedging of variable annuities with state-dependent fees*, Insurance Math. Econom. **58** (2014), 24–33. MR 3257334

[36] G. Di Nunno, B. Øksendal, and F. Proske, *Malliavin calculus for Lévy processes with applications to finance*, Universitext, Springer-Verlag, Berlin, 2009. MR 2460554 (2010f:60001)

[37] R. J. DiPerna and P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98** (1989), no. 3, 511–547. MR 1022305 (90j:34004)

[38] B. Dupire, *Functional Itô Calculus*, Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS. (2009).

[39] N. Eisenbaum, *Integration with respect to local time*, Potential Anal. **13** (2000), no. 4, 303–328. MR 1804175 (2002e:60085)

[40] K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286. MR 1297021 (95j:60087)

[41] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, Characterization and convergence. MR 838085 (88a:60130)

[42] L. C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943 (2011c:35002)

[43] E. Fedrizzi and F. Flandoli, *Pathwise uniqueness and continuous dependence of SDEs with non-regular drift*, Stochastics **83** (2011), no. 3, 241–257. MR 2810591 (2012h:60184)

[44] ______, *Hölder flow and differentiability for SDEs with nonregular drift*, Stoch. Anal. Appl. **31** (2013), no. 4, 708–736. MR 3175794

[45] ______, *Noise prevents singularities in linear transport equations*, J. Funct. Anal. **264** (2013), no. 6, 1329–1354. MR 3017266

[46] F. Flandoli, *Random perturbation of PDEs and fluid dynamic models*, Lecture Notes in Mathematics, vol. 2015, Springer, Heidelberg, 2011, Lectures from the 40th Probability Summer School held in Saint-Flour, 2010. MR 2796837 (2012c:60162)

[47] F. Flandoli, M. Gubinelli, and E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), no. 1, 1–53. MR 2593276 (2011b:35583)

[48] F. Flandoli, T. Nilssen, and P. Proske, *Malliavin differentiability and strong solutions for a class of SDE in Hilbert spaces*, Preprint series. University of Oslo, ISSN 0806-2439 (2013).

[49] E. Fournié, J-M. Lasry, J. Lebuchoux, and P-L. Lions, *Applications of Malliavin calculus to Monte-Carlo methods in finance. II*, Finance Stoch. **5** (2001), no. 2, 201–236. MR 1841717 (2002e:91063)

[50] E. Fournié, J-M. Lasry, J. Lebuchoux, P-L. Lions, and N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch. **3** (1999), no. 4, 391–412. MR 1842285 (2002e:91062)

[51] N. Fournier and J. Printems, *Absolute continuity for some one-dimensional processes*, Bernoulli **16** (2010), no. 2, 343–360. MR 2668905 (2011d:60176)

[52] D. Geman and J. Horowitz, *Occupation densities*, Ann. Probab. **8** (1980), no. 1, 1–67. MR 556414 (81b:60076)

[53] L. Grafakos, *Classical Fourier analysis*, second ed., Graduate Texts in Mathematics, vol. 249, Springer, New York, 2008. MR 2445437 (2011c:42001)

[54] I. Gyöngy and N. Krylov, *Existence of strong solutions for Itô's stochastic equations via approximations*, Probab. Theory Related Fields **105** (1996), no. 2, 143–158. MR 1392450 (97h:60058)

[55] I. Gyöngy and T. Martínez, *On stochastic differential equations with locally unbounded drift*, Czechoslovak Math. J. **51(126)** (2001), no. 4, 763–783. MR 1864041 (2002h:60119)

[56] S. Haadem and F. Proske, *On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients*, J. Funct. Anal. **266** (2014), no. 8, 5321–5359. MR 3177338

[57] M. Hayashi, A. Kohatsu-Higa, and G. Yûki, *Local Hölder continuity property of the densities of solutions of SDEs with singular coefficients*, J. Theoret. Probab. **26** (2013), no. 4, 1117–1134. MR 3119987

[58] T. Hida, H-H. Kuo, J. Potthoff, and L. Streit, *White noise*, Mathematics and its Applications, vol. 253, Kluwer Academic Publishers Group, Dordrecht, 1993, An infinite-dimensional calculus. MR 1244577 (95f:60046)

[59] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. MR 0222474 (36 #5526)

[60] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 1943877 (2003j:60001)

[61] G. Kallianpur and J. Xiong, *Stochastic differential equations in infinite-dimensional spaces*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 26, Institute of Mathematical Statistics, Hayward, CA, 1995, Expanded version of the lectures delivered as part of the 1993 Barrett Lectures at the University of Tennessee, Knoxville, TN, March 25–27, 1993, With a foreword by Balram S. Rajput and Jan Rosinski. MR 1465436 (98h:60001)

[62] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR 1121940 (92h:60127)

[63] C. Klüppelberg, T. Meyer-Brandis, and A. Schmidt, *Electricity spot price modelling with a view towards extreme spike risk*, Quant. Finance **10** (2010), no. 9, 963–974. MR 2738821 (2011j:91263)

[64] A. Kohatsu-Higa and A. Makhlouf, *Estimates for the density of functionals of SDEs with irregular drift*, Stochastic Process. Appl. **123** (2013), no. 5, 1716–1728. MR 3027897

[65] A. Kohatsu-Higa and M. Sanz-Solé, *Existence and regularity of density for solutions to stochastic differential equations with boundary conditions*, Stochastics Stochastics Rep. **60** (1997), no. 1-2, 1–22. MR 1436429 (98d:60111)

[66] N. V. Krylov, *Some new results in the theory of controlled diffusion processes*, Mat. Sb. (N.S.) **109(151)** (1979), no. 1, 146–164, 166. MR 538554 (80j:60104)

[67] ______, *Some properties of traces for stochastic and deterministic parabolic weighted Sobolev spaces*, J. Funct. Anal. **183** (2001), no. 1, 1–41. MR 1837532 (2002c:46069)

[68] N. V. Krylov and M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields **131** (2005), no. 2, 154–196. MR 2117951 (2005k:60209)

[69] A. Kufner, *Weighted Sobolev spaces*, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1985, Translated from the Czech. MR 802206 (86m:46033)

[70] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. 24, Cambridge University Press, Cambridge, 1997, Reprint of the 1990 original. MR 1472487 (98e:60096)

[71] H-H. Kuo, *White noise distribution theory*, Probability and Stochastics Series, CRC Press, Boca Raton, FL, 1996. MR 1387829 (97m:60056)

[72] S. Kusuoka, *Existence of densities of solutions of stochastic differential equations by Malliavin calculus*, J. Funct. Anal. **258** (2010), no. 3, 758–784. MR 2558176 (2011d:60164)

[73] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. I*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306. MR 780762 (86k:60100a)

[74] A. Lanconelli and F. Proske, *On explicit strong solution of Itô-SDE's and the Donsker delta function of a diffusion*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **7** (2004), no. 3, 437–447. MR 2085642 (2005h:60186)

[75] G. Leoni, *A first course in Sobolev spaces*, Graduate Studies in Mathematics, vol. 105, American Mathematical Society, Providence, RI, 2009. MR 2527916 (2010m:46049)

[76] W. V. Li and A. Wei, *A Gaussian inequality for expected absolute products*, J. Theoret. Probab. **25** (2012), no. 1, 92–99. MR 2886380

[77] T. J. Lyons, M. Caruana, and T. Lévy, *Differential equations driven by rough paths*, Lecture Notes in Mathematics, vol. 1908, Springer, Berlin, 2007, Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an introduction concerning the Summer School by Jean Picard. MR 2314753 (2009c:60156)

[78] P. Malliavin, *Stochastic calculus of variation and hypoelliptic operators*, Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), Wiley, New York-Chichester-Brisbane, 1978, pp. 195–263. MR 536013 (81f:60083)

[79] ______, *Stochastic analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 313, Springer-Verlag, Berlin, 1997. MR 1450093 (99b:60073)

[80] B. B. Mandelbrot and J. W. Van Ness, *Fractional Brownian motions, fractional noises and applications*, SIAM Rev. **10** (1968), 422–437. MR 0242239 (39 #3572)

[81] O. Menoukeu-Pamen, T. Meyer-Brandis, T. Nilssen, F. Proske, and T. Zhang, *A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's*, Math. Ann. **357** (2013), no. 2, 761–799. MR 3096525

[82] T. Meyer-Brandis and F. Proske, *On the existence and explicit representability of strong solutions of Lévy noise driven SDE's with irregular coefficients*, Commun. Math. Sci. **4** (2006), no. 1, 129–154. MR 2204081 (2007b:60149)

[83] ______, *Construction of strong solutions of SDE's via Malliavin calculus*, J. Funct. Anal. **258** (2010), no. 11, 3922–3953. MR 2606880 (2011c:60189)

[84] ______, *Explicit representation of strong solutions of SDEs driven by infinite-dimensional Lévy processes*, J. Theoret. Probab. **23** (2010), no. 1, 301–314. MR 2591915 (2011b:60181)

[85] T. Meyer-Brandis and P. Tankov, *Multi-factor jump-diffusion models of electricity prices*, International Journal of Theoretical and Applied Finance **11** (2008), no. 5, 503–528.

[86] K. S. Miller and B. Ross, *An introduction to the fractional calculus and fractional differential equations*, A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1993. MR 1219954 (94e:26013)

[87] S-E. A. Mohammed, T. K. Nilssen, and F. Proske, *Sobolev differentiable stochastic flows for SDEs with singular coefficients: Applications to the transport equation*, Ann. Probab. **43** (2015), no. 3, 1535–1576. MR 3342670

[88] T. Nilssen, *One-dimensional SDE's with discontinuous, unbounded drift and continuously differentiable solutions to the stochastic transport equation*, Preprint series. University of Oslo (2012).

[89] ______, *Quasi-linear stochastic partial differential equations with irregular coefficients-Malliavin regularity of the solutions*, arXiv:1410.4357v2 [math.PR] (2014).

[90] D. Nualart, *The Malliavin calculus and related topics*, second ed., Probability and its Applications (New York), Springer-Verlag, Berlin, 2006. MR 2200233 (2006j:60004)

[91] D. Nualart and Y. Ouknine, *Stochastic differential equations with additive fractional noise and locally unbounded drift*, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, pp. 353–365. MR 2073441 (2005i:60111)

[92] D. Nualart and J. Vives, *Smoothness of Brownian local times and related functionals*, Potential Anal. **1** (1992), no. 3, 257–263. MR 1245230 (94j:60143)

[93] D. Nualart and M. Zakai, *The partial Malliavin calculus*, Séminaire de Probabilités, XXIII, Lecture Notes in Math., vol. 1372, Springer, Berlin, 1989, pp. 362–381. MR 1022924 (90m:60066)

[94] N. Obata, *White noise calculus and Fock space*, Lecture Notes in Mathematics, vol. 1577, Springer-Verlag, Berlin, 1994. MR 1301775 (96e:60061)

[95] B. Øksendal, *Stochastic differential equations*, sixth ed., Universitext, Springer-Verlag, Berlin, 2003, An introduction with applications. MR 2001996 (2004e:60102)

[96] B. Øksendal and A. Sulem, *Applied stochastic control of jump diffusions*, second ed., Universitext, Springer, Berlin, 2007. MR 2322248 (2008b:93003)

[97] L. D. Pitt, *Local times for Gaussian vector fields*, Indiana Univ. Math. J. **27** (1978), no. 2, 309–330. MR 0471055 (57 #10796)

[98] J. Potthoff and L. Streit, *A characterization of Hida distributions*, J. Funct. Anal. **101** (1991), no. 1, 212–229. MR 1132316 (93a:46078)

[99] F. Proske, *Stochastic differential equations—some new ideas*, Stochastics **79** (2007), no. 6, 563–600. MR 2368369 (2008m:60106)

[100] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR 1725357 (2000h:60050)

[101] F. Rezakhanlou, *Regular flows for diffusions with rough drifts*, arXiv:1405.5856 [math.PR] (2014).

[102] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional integrals and derivatives*, Gordon and Breach Science Publishers, Yverdon, 1993, Theory and applications, Edited and with a foreword by S. M. Nikolʹskiĭ, Translated from the 1987 Russian original, Revised by the authors. MR 1347689 (96d:26012)

[103] M. Sanz-Solé, *Malliavin calculus*, Fundamental Sciences, EPFL Press, Lausanne; distributed by CRC Press, Boca Raton, FL, 2005, With applications to stochastic partial differential equations. MR 2167213 (2006h:60005)

[104] A. V. Skorokhod, *Studies in the theory of random processes*, Translated from the Russian by Scripta Technica, Inc, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965. MR 0185620 (32 #3082b)

[105] A. Yu. Veretennikov, *On the strong solutions of stochastic differential equations*, Theory Probab. Appl. **24** (1979), no. 2, 354–366.

[106] I. Vrkoč, *Liouville formula for systems of linear homogeneous Itô stochastic differential equations*, Comment. Math. Univ. Carolinae **19** (1978), no. 1, 141–146. MR 0494493 (58 #13352)

[107] Y. Xue and Y. Xiao, *Fractal and smoothness properties of space-time Gaussian models*, Front. Math. China **6** (2011), no. 6, 1217–1248. MR 2862654 (2012k:62300)

[108] T. Yamada and S. Watanabe, *On the uniqueness of solutions of stochastic differential equations.*, J. Math. Kyoto Univ. **11** (1971), 155–167. MR 0278420 (43 #4150)

[109] L. Yan, *The Euler scheme with irregular coefficients*, Ann. Probab. **30** (2002), no. 3, 1172–1194. MR 1920104 (2003f:60115)

[110] A. K. Zvonkin, *A transformation of the phase space of a diffusion process that will remove the drift*, Mat. Sb. (N.S.) **93(135)** (1974), 129–149, 152. MR 0336813 (49 #1586)